



Published in final edited form as:

*J Alzheimers Dis.* 2021 ; 82(3): 1123–1136. doi:10.3233/JAD-201280.

## Sensitive measures of cognition in Mild Cognitive Impairment

Nathaniel Klooster<sup>a,b,\*</sup>, Stacey Humphries<sup>a,c</sup>, Eileen Cardillo<sup>a,c</sup>, Franziska Hartung<sup>a,c</sup>, Long Xie<sup>d,e</sup>, Sandhitsu Das<sup>d</sup>, Paul Yushkevich<sup>d</sup>, Arun Paliana<sup>d,f</sup>, Jieqiong Wang<sup>a</sup>, David A. Wolk<sup>a,f</sup>, Anjan Chatterjee<sup>a,b,c,f</sup>

<sup>a</sup>Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

<sup>b</sup>Moss Rehabilitation Research Institute, Elkins Park, PA, USA

<sup>c</sup>Penn Center for Neuroaesthetics, University of Pennsylvania, Philadelphia, PA, USA

<sup>d</sup>Penn Image Computing and Science Laboratory (PICSL), University of Pennsylvania, Philadelphia, PA, USA

<sup>e</sup>Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA

<sup>f</sup>Penn Memory Center, University of Pennsylvania, Philadelphia, PA, USA

### Abstract

**Background:** Sensitive measures of cognition are needed in preclinical and prodromal Alzheimer’s disease (AD) to track cognitive change and evaluate potential interventions. Neurofibrillary tangle pathology in AD is first observed in BA35, the medial portion of the perirhinal cortex. The importance of the perirhinal cortex for semantic memory may explain early impairments of semantics in preclinical AD. Additionally, our research has tied figurative language impairment to neurodegenerative disease.

**Objective:** We aim to identify tasks that are sensitive to cognitive impairment in individuals with Mild Cognitive Impairment (MCI), and that are sensitive to atrophy in Brodmann Area 35 (BA35).

**Methods:** Individuals with MCI and cognitively normal participants (CN) were tested on productive and receptive experimental measures of semantic memory and experimental tests of figurative language comprehension (including metaphor and verbal analogy). Performance was related to structural imaging and standard neuropsychological assessment.

**Results:** On the experimental tests of semantics and figurative language, people with MCI performed worse than CN participants. The experimental semantic memory tasks are sensitive and specific; performance on the experimental semantic memory tasks related to MTL structural integrity, including BA35, while standard neuropsychological assessments of semantic memory did not, demonstrating the sensitivity of these experimental measures. A visuo-spatial analogy task did not differentiate groups, confirming the specificity of semantic and figurative language tasks.

\*Correspondence: Nathaniel B. Klooster, Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA 19104, 616 – 994 - 2039, klooster@pennmedicine.upenn.edu.

#### CONFLICT OF INTEREST

Dr. Wolk has received grant funding from Merck, Biogen, and Avid/Eli Lilly; he has received consulting fees from GE Healthcare and Neuronix and is on a DSMB for Functional Neuromodulation.

**Conclusion:** These experimental measures appear sensitive to cognitive change and neurodegeneration early in the AD trajectory and may prove useful in tracking cognitive change in clinical trials aimed at early intervention.

### Keywords

Medial Temporal Lobe; Perirhinal Cortex; Semantic memory; Figurative language; Mild Cognitive Impairment; Alzheimer's disease

---

## INTRODUCTION

Potential interventions in Alzheimer's Disease (AD) are most likely to be effective if implemented as early as possible [1]. As such, sensitive measures of cognition that differentiate the earliest stages of AD from healthy aging are needed [2]. To define this "predementia" period, criteria have been developed to classify individuals as having preclinical and prodromal AD, the latter being the Mild Cognitive Impairment (MCI) stage of disease [3-5].

Cognitive assessments sensitive to preclinical AD are needed. Developing cognitive measures that differentiate preclinical AD from healthy aging promises to improve understanding of the disease and provide a way to track cognitive changes in this population. These measures would help assess the efficacy of interventions when applied early in the disease. While imaging and fluid biomarkers appear sensitive to some aspects of preclinical disease [6-9], a sensitive cognitive measure has greater face validity than biomarkers as outcome measures in clinical trials, since the ultimate goal of treatment is to improve or protect against cognitive decline [10].

The literature offers mixed evidence as to whether standard neuropsychological assessment is sensitive to preclinical AD. Some studies using standard psychometric assessments show promise in differentiating those with preclinical AD from healthy aging [9-11], while others suggest standard measures are relatively insensitive to preclinical AD [12-14].

The current study evaluates the utility of a battery of cognitive tasks predicted to be sensitive to MCI. We hypothesized that measures of semantic richness and comprehension of figurative language are sensitive to cognitive impairment in prodromal AD, and depend on intact MTL subregions, including those first affected in AD. We predicted that people with MCI would perform worse on these tasks than cognitively normal (CN) participants, and that task performance would relate to atrophy in Brodmann Area 35 (BA35).

The pathological processes contributing to AD begin years to decades before there are obvious cognitive consequences as measured with standard tools. Neurofibrillary tangle pathology begins in the medial portion of Brodmann Area 35 (BA35) in the perirhinal cortex, also referred to as the trans-entorhinal region, before spreading medially into entorhinal cortex and the hippocampus, and laterally to cortical temporal regions [15-16]. The perirhinal cortex has been shown to support semantic memory [17-23], especially when making fine-grained semantic distinctions [24-26]. Semantic impairments are reported in preclinical AD [27-29], leading some investigators (e.g. [30]) to propose that semantic

memory should be a greater research focus in these populations rather than the traditional and dominant emphasis on episodic memory.

Standard neuropsychological tests of semantic memory (e.g. picture naming) may not be sensitive enough to detect subtle impairments at the preclinical and prodromal stages of AD. The psycholinguistic and language learning literatures differentiate surface-level knowledge of a word or concept, from deep, rich meaning [31-32]. Shallow word knowledge enables a participant to correctly name a picture or provide a simple 3- or 4-word definition. Deeper mastery involves richer information associated with a word – its different senses, the number of features of a concept, and other words with which it is associated. The N400, an electrophysiological marker of semantic richness, is larger for words with greater number of associates, features, and senses [31].

As an example of the sensitivity of tests of semantic memory, we reported that people with hippocampal amnesia have previously unappreciated deficits of semantic knowledge [33]. Previously acquired, remote semantic knowledge had been considered intact in people with hippocampal amnesia who performed normally on standard clinical neuropsychological assessments of these abilities. However, more sensitive measures reveal that the patients' knowledge was impoverished. While surface level knowledge was intact, patients associated less information with common words than healthy and brain damaged comparison participants. They provided fewer features and senses to target words and identified fewer correctly matching word associates. Similarly, people with preclinical and prodromal AD, who typically have MTL injury, might also be impaired on semantic richness despite normal performance on traditional neuropsychological assessments such as naming and category fluency.

Assessments of figurative language comprehension also show promise as sensitive measures of cognition and may map onto aspects of semantic knowledge. Figurative language, such as metaphor comprehension, can be impaired despite normal literal language abilities [34-35]. Successful metaphor processing requires working memory, cognitive flexibility, inhibition of literal meaning, abstract thinking, executive demands, and semantic memory. Decrements in any of these subdomains can lead to impairment. Metaphor comprehension is impaired following left-hemisphere neurodegeneration [36], and in AD [37-39]. Figurative language abilities more broadly are impaired in AD [40] and Mild Cognitive Impairment [41]. Metaphor comprehension may be an especially fragile cognitive ability and sensitive to cognitive decline early in the AD trajectory.

The current study evaluates the usefulness of a battery of tasks predicted to be sensitive to early stages of AD in a sample of cognitively normal older adults and people with MCI, a population enriched in individuals with prodromal AD. We predicted that tests of semantic richness would better discriminate the groups than standard neuropsychological assessments of semantic memory in prodromal AD. Tests of metaphor comprehension and verbal analogy were chosen because the complex cognitive skills needed to successfully resolve the meaning of novel figurative language may be specifically affected in early AD. A visuo-spatial analogy task was included as a control task to evaluate a potential dissociation from language and semantic tasks. We also examined brain-behavior relationships in this

population. Our primary anatomical analyses relate task performance to integrity of BA35, the first region affected by neurofibrillary tangles of AD. If these tasks rely on the integrity of this region, this brain-behavior relationship would support a potential role for these measures in preclinical AD. Finally, we conducted exploratory analyses to examine relationships with other MTL subregions and with brain regions outside the MTL.

## MATERIALS AND METHODS

### Participants

Cognitively normal older adults (CN) included 56 participants (35 female) without prior medical or neurologic conditions that may affect cognition. CN participants were an average of 72.84 ( $\pm 7.34$ ) years old and completed an average of 15.53 ( $\pm 2.75$ ) years of education. CN participants include 30 Caucasian participants, 22 African American, 2 multi-racial, and 2 Latinx participants. All CN (MOCA:  $M = 27.10 \pm 2.25$ ) participants are enrolled in the Penn Alzheimer's Disease Core Center where they undergo annual medical, neurological and psychometric evaluation. All CN participants have a Clinical Dementia Rating score of 0. Consensus designation of cognitively normal status is made annually.

Participants with a clinical diagnosis of amnesic Mild Cognitive Impairment (MCI) consisted of 35 people (17 female). Two participants were borderline MCI; They did not have a subjective complaint, but had objective evidence of cognitive decline, and within the subsequent year were diagnosed with amnesic MCI. MCI status was determined based on clinical evaluation and standard psychometric testing following NIA-AA guidelines [4]. This group does not include individuals with prior neurological or medical conditions that affect cognition. Patients were matched to the CN group on age ( $72.68 \pm 6.84$ ,  $p > 0.9$ ), but completed more years of formal education on average ( $17.06 \pm 2.57$ ,  $p < 0.01$ ). MCI patients include 28 Caucasian participants and 7 African American participants. All participants completed evaluations from the Uniform Data Set (UDS 3: [42]). MCI patients performed worse than CN participants on tests of overall cognition (MOCA  $p < 0.01$ ) indicating mildly impaired cognition, and on tests of semantic memory (Animal fluency:  $p < 0.01$ ; Vegetable fluency:  $p < 0.01$ ) and other domains (see Table 1).

All participants were native English speakers, gave informed consent in accordance with procedures of the University of Pennsylvania Institutional Review Board, and were paid \$20/hour for their time. The research was conducted in accordance with the Declaration of Helsinki.

### Experimental tasks

**Senses listing task**—Our productive measure of semantic richness (See [33] for task details), the senses task, presents participants with a target word (e.g. Pen.). Participants are given one minute to list as many different senses of the word as possible (e.g. “A writing instrument,” “To write a letter,” “An enclosure for animals,” “University of Pennsylvania”). Stimuli include the same 20 words chosen from normed databases [43-44]. Responses are recorded for later transcription and analysis.

## Transcription and coding of responses.

The first author transcribed all responses. A research assistant, blind to participant status and study hypotheses, transcribed 10% of the data for each participant chosen at random. This transcription was compared to the first author's transcription to examine agreement. Interrater reliability was  $\kappa = 0.89$ , indicating near perfect agreement [45]. These transcriptions were then evaluated and coded. Responses were evaluated with those in Wordnet, a normative database for polysemy [46], and with the Merriam Webster Dictionary [47]. Responses included in Wordnet were coded as correct responses, as were responses matching one of the definitions in Merriam Webster. Homophones to the target word and proper nouns were coded as correct responses. Other responses were coded as incorrect and were excluded from analysis.

**Word Associates Task (WAT)**—For a receptive measure of depth of semantic knowledge, participants are given a paper and pencil version of John Read's 1998 WAT form [32]. Consisting of 40 items, each has a target word in bold with eight possible associates below it and participants choose the four correctly matching associates. Words on the left are possible synonyms and words on the right are possible collocates, words that might follow the target word in a phrase or a sentence. After instructions explaining the task, participants complete two practice problems and receive feedback. Each correct response is scored for a total of 160 possible points. Points are not deducted for incorrect responses.

## Metaphor

**Stimuli:** Stimuli from the Metaphor Multiple Choice task [35, 48-49] include 120 sentences, 60 novel metaphorical sentences, and 60 literal sentences, matched on the base term of the metaphor (See Table 2). The 60 sentence-pairs for this study were selected from a normative database designed for neuropsychological and imaging studies [49, 50] with the aid of Stochastic Optimization of Stimuli [51]. Sentences were matched group-wise on average frequency and concreteness, familiarity, naturalness, imageability, number of words, and number of content words.

**Answer choices:** Each sentence has four possible answers, the correct target answer and three foils. Answers consist of two words, an adjective-noun combination. Foils for literal sentences include 1) a category associate of the agent of the sentence not implied by the sentence, 2) the opposite of the literal meaning, and 3) an unrelated answer. Foils for the metaphor stimuli include 1) the literal meaning of the sentence, 2) the opposite of the metaphorical meaning, and 3) an unrelated answer. Answers were matched on average frequency and concreteness.

**Procedure:** The metaphor task was run in E-Prime 2.0. The target sentence appeared at the top of the screen, with the four answers in a square below it. Participants were instructed to select the answer that best expresses the meaning of the sentence. The position of the target and foils were randomized per item. Participants responded using keyboard button presses, responding as quickly and accurately as possible. The order of sentences was randomized across participants. Instructions were read aloud, and participants completed practice trials to ensure comprehension of the task

**Verbal Analogy:** Participants completed the verbal analogy task based on Green et al., [52]. As depicted in Figure 1, participants see a relationship at the top of the screen and two possible response sets at the bottom. Participants choose one set from the possibilities at the bottom that are related in the same way as the words at the top. The task was administered through ePrime and included 30 trials. Order of trials was randomized across participants. Each correct response was scored for a total of 30 possible points.

**Visual-spatial analogy:** The visual analogy task served as a control task. Participants completed the visual analogy task described in Watson & Chatterjee [53]. As depicted in Figure 1, for each trial three sets of colored shapes were presented, the source set at the top of the screen and two possible response sets below it. Participants chose the response set that contained the same pattern of spatial relations as the source. The pattern can be based on color or shape. When the basis of the analogy was the relations between colors, all sets contained the same pattern of shapes. When the basis of the analogy was the relations of shapes, all sets contained the same pattern of colors. The task was administered through ePrime 2.0 and included 30 trials. Order of trials was randomized across participants. Each correct response was scored for a total of 30 possible points.

## Statistical Methods

For behavioral group analyses, Generalized Linear Mixed-Effects Modelling (GLMEM) was used to predict group membership (CN or MCI) from task performance with random effects of item and participant. For anatomical analyses, Linear Mixed-Effects Modelling (LMEM) was used to predict task performance from fixed effects of anatomy, age, and education and random effects of item and participant. The lme4 package [54] in R (Version 3.6.3) was used for the following analyses.

Linear discriminant analysis was used to compare the relative performance of cognitive variables in discriminating participant group (CN and MCI). Validation procedures were not used, so these analyses are considered exploratory. Standard neuropsychological assessments (all tests listed in Table 1) were compared to the experimental measures (Senses, WAT, Literal trials, Metaphor trials, Verbal Analogy). Table 7 lists all the variable included in this analysis. Tests with higher coefficients of linear discrimination indicate a greater ability to differentiate groups. The *lda* function from the MASS package in R (Version 3.6.3) was used for these analyses.

## Anatomical Methods

MRI scans were acquired on a 3T Siemens Prisma scanner at the Hospital of the University of Pennsylvania using a 64-channel array coil. The protocol includes T1-weighted (MPRAGE) whole-brain scan with the following parameters: TR/TE/TI=2400/2.24/1060 ms, 8° flip angle,  $0.8 \times 0.8 \times 0.8$  mm<sup>3</sup> resolution, acquisition time 6:38 min. MTL segmentation was performed in T1-weighted MRI using an analysis pipeline optimized for extracting subregional MTL measures, including anterior/posterior hippocampus, entorhinal cortex (ERC), BA35 (i.e. transentorhinal cortex), BA36, and parahippocampal cortex using an automatic pipeline, Automatic Segmentation of Hippocampal Subfields – T1 (ASHS-T1; [55]). This technique is tailored to reduce confounds of other approaches in this region, such

as anatomic variability and segmentation of dural tissue from grey matter [55]. Anterior/posterior hippocampal volumes and cortical thickness of extrahippocampal MTL subregions (entorhinal cortex, BA35, BA36 and parahippocampal cortex) were calculated from ASHS-T1 output [56]. See supplemental materials for additional details.

The participant groups were combined to examine brain-behavior relationships. LMEMs predicted task performance from the fixed effects of group, age and education and the random effects of participant and task item. First, the relationship between integrity of our region of interest, BA35, and task performance was examined. Next, exploratory analyses were run examining integrity of other MTL subregions and performance. The Holm method was used to control for multiple comparisons.

Finally, exploratory whole-brain cortical thickness analysis was performed using the ANTs cortical thickness pipeline that implements the DiReCT thickness estimation method [57]. These analyses relate performance on each task to cortical thickness in the collapsed groups. Regression analyses between cortical thickness values and performance on each task was performed at each voxel in a population template space. The randomize tool in FSL was used to run non-parametric permutations ( $n = 10,000$ ) for each score. Clusters that met a height threshold of  $p < 0.05$  uncorrected with threshold-free cluster enhancement and a minimum of 25 adjacent voxels are reported.

## RESULTS

Tests of semantic memory, metaphor comprehension, and analogy comprehension were evaluated for their sensitivity to MCI and to cortical thickness of BA35.

### Behavioral Results

**Senses Task**—GLMEM analyses (Table 3 and Figure 2A) revealed a significant effect of Senses performance ( $p < 0.05$ ) in differentiating groups, with CN outperforming MCI.

**WAT**—GLMEM analyses (Table 4 and Figure 2B) reveal a significant effect of WAT performance ( $p < 0.005$ ) in differentiating groups, with CN participants performing better than MCI.

**Metaphor Task**—Separate GLMEMs were used to analyze the effect of task performance on metaphor and literal trials in differentiating groups. These analyses reveal that Literal trial performance differentiates the CN and MCI groups ( $p < 0.01$ ) while metaphor trials do not ( $p > 0.08$ , Table 5 and Figure 2C).

**Verbal Analogy**—GLMEM analyses (Table 6 and Figure 2D) reveal task performance significantly differentiates groups ( $p < 0.05$ ) with CN outperforming MCI.

**Visual-spatial Analogy**—To examine potential group differences, 52 CN participants and 35 people with MCI completed the Visual-spatial Analogy Task. GLMEM analyses show that visuo-spatial analogy task performance did not differentiate groups ( $p > 0.1$ ). There

does not appear to be floor or ceiling effects as both groups perform significantly above chance ( $p < 0.05$ ) and perform more than 2 standard deviations away from a perfect score.

**Linear discriminant analyses**—Exploratory linear discriminant analyses were performed to examine which experimental and neuropsychological tests most strongly differentiate groups (MCI vs CN, Table 7). Tests with higher discriminant coefficients more strongly differentiate groups. These analyses reveal that the experimental measures better differentiate MCI from CN groups than standard neuropsychological assessments.

In summary, people with MCI perform worse on the tests of semantic richness, on the metaphor task, and on the verbal analogy task. Groups do not differ on the visuo-spatial analogy task. The experimental tasks appear to better differentiate groups than the standard neuropsychological assessments.

### Anatomical results

**Senses Task**—Anatomical data was available from 25 participants with MCI who completed the senses task and 52 CN participants.

First, the relationship between integrity of our region of interest, BA35, and task performance was examined. Across groups, mean BA35 cortical thickness was associated with senses task performance ( $\beta = 0.77$ ,  $p < 0.001$ , Table 8 and Figure 3). A 1 mm increase in BA35 thickness is associated with a 0.77 increase in the number of senses produced. Within the CN group, mean BA35 thickness was associated with the number of senses produced ( $p < 0.05$ ). Within the MCI group, no relationships were seen between BA35 integrity and task performance

Next, exploratory analyses were run examining integrity of other MTL subregions and task performance.

Across the combined groups, significant relationships were seen between senses task performance and integrity of mean anterior hippocampal volume ( $p < 0.05$ ), mean BA36 thickness ( $p < 0.001$ ) and mean parahippocampal cortex thickness ( $p < 0.005$ ).

Next, we examined relationships between MTL structural integrity and task performance in each group separately. Within the MCI group, no relationships were seen between MTL subregional integrity and task performance. Within the CN group, mean parahippocampal cortex thickness ( $p < 0.01$ ) was associated with senses produced.

Finally, exploratory full-brain analyses revealed no relationship between cortical thickness and task performance within or outside the MTL.

**WAT**—Anatomical data was available from 26 participants with MCI who completed the WAT and 55 CN participants.

Across groups, mean BA35 thickness was associated with WAT performance ( $\beta = 0.191$ ,  $p < 0.05$ , Table 9 and Figure 4). Left BA35 thickness ( $\beta = 0.287$ ) was more strongly associated with task performance than right BA35 ( $\beta = 0.199$ ). Within the MCI group, no relationship



is seen between BA35 thickness and task performance ( $p > 0.1$ ). Within the CN group, mean BA35 thickness was associated with WAT performance ( $p < 0.05$ ).

Across groups, no other MTL subregions show a relationship with task performance. Within the MCI group, no relationships were seen between MTL subregional integrity and WAT performance. Within the CN group, mean parahippocampal cortex thickness was associated with performance ( $p < 0.001$ ).

Exploratory full-brain voxel-based morphometry analyses reveal relationships between WAT performance and atrophy in the left hippocampus, the left anterior parahippocampal gyrus, and the right inferior temporal gyrus (Table 10 and Figure 5).

**Metaphor**—Anatomical data was available from 25 participants with MCI who completed the metaphor task and 51 CN participants.

For metaphor trials, mean BA35 thickness showed a trend in predicting accuracy ( $p < 0.1$ ) across groups. Among CN participants, mean BA35 thickness showed a trend in predicting metaphor accuracy ( $p < 0.06$ ). There was no relationship between BA35 and MCI task performance.

No other MTL subregions showed a relationship with metaphor accuracy across groups, or within the MCI or CN groups.

For literal trials, mean parahippocampal cortex thickness was associated with accuracy ( $p < 0.05$ ) across groups. Within both the MCI and CN participant groups, no MTL subregions predicted literal accuracy. Exploratory full-brain analyses revealed no relationship between cortical thickness and task performance outside the MTL.

**Analogy Tasks**—Anatomical data was available from 52 CN participants and 28 participants with MCI who completed the verbal analogy and visual-spatial analogy tasks. There was no relationship between MTL subregional integrity and performance on either task.

Exploratory full-brain analyses reveal no relationship between cortical thickness and task performance outside the MTL.

**Neuropsychological assessment and anatomy**—MCI patients (Table 1) performed worse than CN participants on tests of overall cognition (MOCA  $p < 0.01$ ) and on tests of semantic memory (Animal fluency  $p < 0.01$ . Vegetable fluency  $p < 0.01$ ).

Animal ( $p > 0.9$ ) and vegetable ( $p > 0.6$ ) fluency did not relate to BA35 thickness, or to other MTL subregion integrity ( $ps > 0.2$ ).

In summary for the anatomical analyses, performance on the semantic richness tasks showed relationships with BA35 integrity across groups and within the CN group. Metaphor task performance showed a trend. Performance on the analogy tasks and on standard neuropsychological assessments did not relate to atrophy in BA35.

## DISCUSSION

The goal of this study was to determine if experimental measures of semantic knowledge and figurative language can discriminate MCI from cognitively normal older adults, and to determine the degree to which performance on these tasks relates to morphometric measures of the earliest regions of AD pathology. Such measures, if sensitive to cognitive decline in the earliest stages of Alzheimer's disease (AD), promise to help our ability to monitor disease progression and response to treatment. We tested the hypothesis that semantic richness is impaired in early stages of AD because of the recently appreciated role of the medial temporal lobe in supporting semantics, and the fact that these areas are known to be affected early in the disease. We also tested the hypothesis that comprehension of figurative language is impaired in this population. As predicted, these measures differentiated people with MCI from CN participants, and appear better at differentiating groups than standard psychometric tests. Further, performance on the experimental semantic measures, but not the figurative language tests, nor the standard neuropsychological assessments, was related to integrity of BA35, which includes the transentorhinal region, the first region associated with the neurofibrillary tangle pathology of AD.

On productive and receptive measures of semantic richness and depth of knowledge, MCI patients performed worse than CN participants and performance related to integrity of MTL subregions. Previous work has documented impairments on these tasks in patients with focal and stable MTL lesions [33, 58]. In the current study, semantic performance across groups, and remarkably within the CN group, related to integrity of BA35. Previous research has demonstrated a role for perirhinal cortex, including BA35, in supporting rich semantic memory [24-26]. It is striking that the standard neuropsychological assessments of semantic memory, which contributed to diagnostic classification, did not differentiate the groups with as large of an effect size, nor relate as well to MTL atrophy, supporting the claim of greater sensitivity for our experimental measures.

Given the involvement of MTL structures with early AD pathology, studies of episodic memory have been a focus of the study of prodromal [59-61] and preclinical AD [62-64]. While episodic memory is undoubtedly affected early in the course of AD, language is another aspect of cognition that notably deteriorates with disease progression [65-67]. Further, the perirhinal cortex, the area first affected by AD pathology, has long been tied to semantic memory [17, 20, 22, 29]. The current results argue that greater attention should be focused on semantic memory and figurative language in the study of preclinical and prodromal AD.

The ability to understand novel metaphorical and literal sentences differentiates CN from MCI participants. Previous work has documented metaphor impairments in the absence of literal language difficulty (as assessed by standard psychometric tools), in patients with focal brain lesions [35] and patients with neurodegenerative disease [36, 48]. These studies show that metaphor processing can be disproportionately impaired compared to literal sentence comprehension, suggesting that metaphor comprehension deficits could be a sensitive measure of cognitive change, revealing impairments while literal language is normal. Understanding a novel metaphor is a complex cognitive achievement requiring

inhibition of literal meaning, and cognitive flexibility to compare concepts from different domains to highlight their shared features. Impairments in multiple aspects of cognition may lead to a metaphor comprehension impairment. With its distributed neural support and cognitive complexity, metaphor comprehension may be sensitive to early cognitive changes in a variety of neurodegenerative disease.

MCI participants showed deficits on the semantic tasks, the metaphor task, and the verbal analogy task. The lack of group difference on the visuo-spatial analogy task, which also requires complex reasoning abilities, demonstrates the specificity of these findings. It is not the case that MCI patients are impaired at any complex task; their impairment in our study was restricted to tests that involved language and meaning.

Importantly, relationships were observed between task performance and integrity of MTL subregions not only across groups, but also within the CN group. The CN group likely contains individuals with preclinical AD and varying degrees of changes to MTL subregional integrity. Our morphometric measure of BA35 has been shown to be sensitive to the presence of preclinical AD in cognitively normal populations [68]. The present results suggest these cognitive tasks are related to these subtle changes and, thus, provide promise that they might have utility in detecting cognitive consequences of preclinical AD. It is unclear why this relationship is absent in the MCI group. Perhaps degeneration beyond the MTL diminishes the specific relationship between BA35 and task performance. Additionally, the smaller sample size of this group may have reduced power to detect such a relationship.

In our exploratory analyses we confirmed the relationship between hippocampal integrity and performance on the semantic memory tasks, revealing the necessity of the medial temporal lobe (MTL) for semantic richness, replicating previous findings [33]. These exploratory analyses failed to uncover relationships between task performance and neural regions outside of the medial temporal lobes, suggesting that MTL injury is sufficient to modulate performance on these tasks. This lack of correlation with additional cortical regions also may relate to that fact that in this population, tangle pathology, if present, would be largely confined to MTL structures. We predict that later in the course of AD, when degeneration spreads beyond the MTL, we would see relationships between degeneration and performance on the figurative language tasks. Novel metaphor comprehension, as tested in the current study, has previously been linked to neural activity in bilateral inferior frontal gyri and the left posterior middle temporal gyrus [36, 69]. While the CN group outperformed MCI on the Verbal Analogy task, performance did not relate to MTL subregional anatomy. Task performance was previously tied to frontopolar activity [52]. Groups did not differ on Visuospatial Analogy performance and no relationship was observed with MTL integrity. Visuospatial analogy task performance was previously linked to bilateral rostrolateral prefrontal cortex [53]. In more advanced AD, when degeneration spreads to the frontal lobes, there would likely be stronger relationships seen with figurative language tasks.

A limitation of the current study is not having molecular markers of AD in enough individuals to include in the current analysis. Some participants in the CN group likely had

preclinical AD with evidence of cerebral amyloid. Accounting for such participants would likely further differentiate preclinical and prodromal AD from healthy aging. Conversely, some MCI participants would likely show an absence of cerebral amyloid and have cognitive impairment because of another etiology. Accounting for these individuals would give a clearer view of cognitive impairment specifically due to prodromal AD.

Other limitations include the relatively high level of education completed by our participants. The MCI group had a significantly higher level of education than the CN group, so we would expect to see even greater group differences if the groups were matched. Education was controlled for in all statistical models. That said, it remains to be seen how groups with lower levels of education would perform on these experimental measures. Finally, the current findings apply to native English speakers only. How non-native speakers would perform on these tasks is unknown.

Future directions for investigation include relating task performance to both amyloid and tau PET imaging. Useful cognitive screening instruments would ideally, in those with normal cognition, differentiate participants with and without cerebral amyloid. As the presence and location of neurofibrillary tangles often mirror the pattern of neurodegeneration and resulting cognitive decline [70-71], we predict that semantic task performances will be especially sensitive to the presence of Tau.

In summary, MCI patients performed worse than CN participants on measures of semantic richness and figurative language. Experimental semantic task performance relates to atrophy of BA35, the area first impacted by AD pathology, while standard neuropsychological assessments of semantics do not. The cognitive instruments evaluated here show promise in sensitivity to early stages of clinical AD.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS:

Funding sources:

The Dolores Smith Innovation Fund

NK supported by T32 HD071844

AG055005, AG010124

## CITATIONS

- [1]. Brooks LG, Loewenstein DA (2010) Assessing the progression of mild cognitive impairment to Alzheimer's disease: current trends and future directions. *Alz Res Therapy* 2, 28.
- [2]. Mortamais M, Ash JA, Harrison J, Kaye J, Kramer J, Randolph C, Pose C, Alcala B, Ropacki M, Ritchie CW, Ritchie K (2017) Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimer's & Dementia* 13, 468–492.
- [3]. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo

MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH (2011) 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* 7, 280–292.

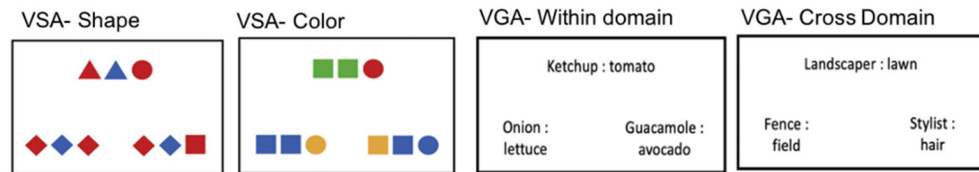
- [4]. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Gamst A, Holtzman DM, Jagust WJ, Petersen RC, Snyder PJ, Carrillo MC, Thies B, Phelps CH (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia* 7, 270–279.
- [5]. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemers E, Snyder HM, Sperling R, Elliott C, Masliah E, Ryan L, Silverberg N (2018) NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia* 14, 535–562.
- [6]. Sutphen CL, Jasielc MS, Shah AR, Macy EM, Xiong C, Vlassenko AG, Benzinger TLS, Stoops EEJ, Vanderstichele HMJ, Brix B, Darby HD, Vandijck MLJ, Ladenson JH, Morris JC, Holtzman DM, Fagan AM (2015) Longitudinal Cerebrospinal Fluid Biomarker Changes in Preclinical Alzheimer Disease During Middle Age. *JAMA Neurol* 72, 1029. [PubMed: 26147946]
- [7]. Villemagne VL, Pike KE, Chételat G, Ellis KA, Mulligan RS, Bourgeat P, Ackermann U, Jones G, Szoek C, Salvado O, Martins R, O'Keefe G, Mathis CA, Klunk WE, Ames D, Masters CL, Rowe CC (2011) Longitudinal assessment of A $\beta$  and cognition in aging and Alzheimer disease. *Ann Neurol* 69, 181–192. [PubMed: 21280088]
- [8]. Vellas B, Carrillo MC, Sampaio C, Brashear HR, Siemers E, Hampel H, Schneider LS, Weiner M, Doody R, Khachaturian Z, Cedarbaum J, Grundman M, Broich K, Giacobini E, Dubois B, Sperling R, Wilcock GK, Fox N, Scheltens P, Touchon J, Hendrix S, Andrieu S, Aisen P, EU/US/CTAD Task Force Members (2013) Designing drug trials for Alzheimer's disease: What we have learned from the release of the phase III antibody trials: A report from the EU/US/CTAD Task Force. *Alzheimer's & Dementia* 9, 438–444.
- [9]. Bondi MW, Salmon DP, Galasko D, Thomas RG, Thal LJ (1999) Neuropsychological function and apolipoprotein E genotype in the preclinical detection of Alzheimer's disease. *Psychology and Aging* 14, 295–303. [PubMed: 10403716]
- [10]. Bondi MW, Jak AJ, Delano-Wood L, Jacobson MW, Delis DC, Salmon DP (2008) Neuropsychological Contributions to the Early Identification of Alzheimer's Disease. *Neuropsychol Rev* 18, 73–90. [PubMed: 18347989]
- [11]. Alzheimer's Disease Neuroimaging Initiative, Thomas KR, Edmonds EC, Eppig J, Salmon DP, Bondi MW (2018) Using Neuropsychological Process Scores to Identify Subtle Cognitive Decline and Predict Progression to Mild Cognitive Impairment. *JAD* 64, 195–204. [PubMed: 29865077]
- [12]. Rentz DM, Parra Rodriguez MA, Amariglio R, Stern Y, Sperling R, Ferris S (2013) Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: a selective review. *Alzheimers Res Ther* 5, 58. [PubMed: 24257331]
- [13]. Loewenstein DA, Curiel RE, Duara R, Buschke H (2018) Novel Cognitive Paradigms for the Detection of Memory Impairment in Preclinical Alzheimer's Disease. *Assessment* 25, 348–359. [PubMed: 29214859]
- [14]. Xie L, Wisse LEM, Das SR, Vergnet N, Dong M, Ittyerah R, Flores R, Yushkevich PA, Wolk DA, for the Alzheimer's Disease Neuroimaging Initiative (2020) Longitudinal atrophy in early Braak regions in preclinical Alzheimer's disease. *Hum Brain Mapp* 41, 4704–4717. [PubMed: 32845545]
- [15]. Braak H, Braak E (1991) Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* 82, 239–259. [PubMed: 1759558]
- [16]. Taylor KI, Probst A (2008) Anatomic localization of the transentorhinal region of the perirhinal cortex. *Neurobiology of Aging* 29, 1591–1596. [PubMed: 17478012]
- [17]. Venneri A, McGeown WJ, Hietanen HM, Guerrini C, Ellis AW, Shanks MF (2008) The anatomical bases of semantic retrieval deficits in early Alzheimer's disease. *Neuropsychologia* 46, 497–510. [PubMed: 17936858]

- [18]. Wang W-C, Lazzara MM, Ranganath C, Knight RT, Yonelinas AP (2010) The Medial Temporal Lobe Supports Conceptual Implicit Memory. *Neuron* 68, 835–842. [PubMed: 21144998]
- [19]. Davies RR, Graham KS, Xuereb JH, Williams GB, Hodges JR (2004) The human perirhinal cortex and semantic memory: Human perirhinal cortex and semantic memory. *European Journal of Neuroscience* 20, 2441–2446.
- [20]. Taylor KI, Moss HE, Stamatakis EA, Tyler LK (2006) Binding crossmodal object features in perirhinal cortex. *Proceedings of the National Academy of Sciences* 103, 8239–8244.
- [21]. Murray EA, Richmond BJ (2001) Role of perirhinal cortex in object perception, memory, and associations. *Current Opinion in Neurobiology* 11, 188–193. [PubMed: 11301238]
- [22]. Clarke A, Tyler LK (2014) Object-Specific Semantic Coding in Human Perirhinal Cortex. *J Neurosci* 34, 4766–4775. [PubMed: 24695697]
- [23]. Hirni DI, Kivisaari SL, Monsch AU, Taylor KI (2013) Distinct neuroanatomical bases of episodic and semantic memory performance in Alzheimer’s disease. *Neuropsychologia* 51, 930–937. [PubMed: 23369803]
- [24]. Tyler LK, Stamatakis EA, Bright P, Acres K, Abdallah S, Rodd JM, Moss HE (2004) Processing Objects at Different Levels of Specificity. *Journal of Cognitive Neuroscience* 16, 351–362. [PubMed: 15072671]
- [25]. Kivisaari SL, Tyler LK, Monsch AU, Taylor KI (2012) Medial perirhinal cortex disambiguates confusable objects. *Brain* 135, 3757–3769. [PubMed: 23250887]
- [26]. Wright P, Randall B, Clarke A, Tyler LK (2015) The perirhinal cortex and conceptual processing: Effects of feature-based statistics following damage to the anterior temporal lobes. *Neuropsychologia* 76, 192–207. [PubMed: 25637774]
- [27]. Papp KV, Rentz DM, Orlovsky I, Sperling RA, Mormino EC (2017) Optimizing the preclinical Alzheimer’s cognitive composite with semantic processing: The PACC5. *Alzheimer’s & Dementia: Translational Research & Clinical Interventions* 3, 668–677.
- [28]. Vonk MJ, Flores RJ, Rosado D, Qian C, Cabo R, Habegger J, Louie K, Allocco E, Brickman AM, Manly JJ (2019) Semantic network function captured by word frequency in nondemented APOE e4 carriers. *Neuropsychology* 33, 256–262. [PubMed: 30489116]
- [29]. Hirni DI, Kivisaari SL, Krumm S, Monsch AU, Berres M, Oeksuez F, Reinhardt J, Ulmer S, Kressig RW, Stippich C, Taylor KI (2016) Neuropsychological Markers of Medial Perirhinal and Entorhinal Cortex Functioning are Impaired Twelve Years Preceding Diagnosis of Alzheimer’s Dementia. *JAD* 52, 573–580. [PubMed: 27031465]
- [30]. Venneri A, Mitolo M, De Marco M (2016) Paradigm shift: semantic memory decline as a biomarker of preclinical Alzheimer’s disease. *Biomarkers Med* 10, 5–8.
- [31]. Pexman PM, Hargreaves IS, Siakaluk PD, Bodner GE, Pope J (2008) There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review* 15, 161–167. [PubMed: 18605497]
- [32]. Read J (2013) Validating a test to measure depth of vocabulary knowledge. In *Validation in language assessment*, Routledge, pp.55 – 74.
- [33]. Klooster NB, Duff MC (2015) Remote semantic memory is impoverished in hippocampal amnesia. *Neuropsychologia* 79, 42–52. [PubMed: 26474741]
- [34]. Cardillo ER, McQuire M, Chatterjee A (2018) Selective Metaphor Impairments After Left, Not Right, Hemisphere Injury. *Front Psychol* 9, 2308. [PubMed: 30559690]
- [35]. Ianni GR, Cardillo ER, McQuire M, Chatterjee A (2014) Flying under the radar: figurative language impairments in focal lesion patients. *Front Hum Neurosci* 8, 871. [PubMed: 25404906]
- [36]. Klooster N, McQuire M, Grossman M, McMillan C, Chatterjee A, Cardillo E (2020) The Neural Basis of Metaphor Comprehension: Evidence from Left Hemisphere Degeneration. *Neurobiology of Language* 1, 474–491.
- [37]. Winner E, Gardner H (1977) The comprehension of metaphor in brain-damaged patients. *Brain* 100, 717–729. [PubMed: 608117]
- [38]. Papagno C (2001) Comprehension of metaphors and idioms in patients with Alzheimer’s disease: A longitudinal study. *Brain* 124, 1450–1460. [PubMed: 11408339]
- [39]. Amanzio M, Geminiani G, Leotta D, Cappa S (2008) Metaphor comprehension in Alzheimer’s disease: Novelty matters. *Brain and Language* 107, 1–10. [PubMed: 17897706]

- [40]. Papagno C, Lucchelli F, Muggia S, Rizzo S (2003) Idiom comprehension in Alzheimer's disease: the role of the central executive. *Brain* 126, 2419–2430. [PubMed: 12902312]
- [41]. Cardoso S, Silva D, Maroco J, de Mendonça A, Guerreiro M (2014) Non-literal language deficits in mild cognitive impairment: Non-literal language deficits in MCI. *Psychogeriatrics* 14, 222–228. [PubMed: 25495083]
- [42]. Weintraub S, Salmon D, Mercaldo N, Ferris S, Graff-Radford NR, Chui H, Cummings J, DeCarli C, Foster NL, Galasko D, Peskind E, Dietrich W, Beekly DL, Kukull WA, Morris JC (2009) The Alzheimer's Disease Centers' Uniform Data Set (UDS): The Neuropsychologic Test Battery. *Alzheimer Disease & Associated Disorders* 23, 91–101. [PubMed: 19474567]
- [43]. Nelson DL, McEvoy CL, Walling JR, Wheeler JW (1980) The University of South Florida homograph norms. *Behavior Research Methods & Instrumentation* 12, 16–37.
- [44]. Twilley LC, Dixon P, Taylor D, Clark K (1994) University of Alberta norms of relative meaning frequency for 566 homographs. *Mem Cogn* 22, 111–126.
- [45]. Landis JR, Koch GG (1977) The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 159. [PubMed: 843571]
- [46]. Miller GA (1995) WordNet: a lexical database for English. *Commun ACM* 38, 39–41.
- [47]. Merriam-Webster, Inc, ed. (2016) The Merriam-Webster dictionary, Merriam-Webster, Incorporated, Springfield, Massachusetts. On-line at <http://www.mw.com/home.htm>.
- [48]. Humphries S, Klooster N, Cardillo E, Weintraub D, Rick J, Chatterjee A (2019) From action to abstraction: The sensorimotor grounding of metaphor in Parkinson's disease. *Cortex* 121, 362–384. [PubMed: 31678683]
- [49]. Cardillo ER, Schmidt GL, Kranjec A, Chatterjee A (2010) Stimulus design is an obstacle course: 560 matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behavior Research Methods* 42, 651–664. [PubMed: 20805587]
- [50]. Cardillo ER, Watson C, Chatterjee A (2017) Stimulus needs are a moving target: 240 additional matched literal and metaphorical sentences for testing neural hypotheses about metaphor. *Behav Res* 49, 471–483.
- [51]. Armstrong BC, Watson CE, Plaut DC (2012) SOS! An algorithm and software for the stochastic optimization of stimuli. *Behav Res* 44, 675–705.
- [52]. Green AE, Kraemer DJM, Fugelsang JA, Gray JR, Dunbar KN (2010) Connecting Long Distance: Semantic Distance in Analogical Reasoning Modulates Frontopolar Cortex Activity. *Cerebral Cortex* 20, 70–76. [PubMed: 19383937]
- [53]. Watson CE, Chatterjee A (2012) A bilateral frontoparietal network underlies visuospatial analogical reasoning. *NeuroImage* 59, 2831–2838. [PubMed: 21982934]
- [54]. Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *J Stat Soft* 67, 13 – 27.
- [55]. Xie L, Wisse LE, Pluta J, de Flores R, Piskin V, Manjón JV, Wang H, Das HR, Ding SL, Wolk DA, & Yushkevich PA (2019). Automated segmentation of medial temporal lobe subregions on in vivo T1-weighted MRI in early stages of Alzheimer's disease. *Human brain mapping*.
- [56]. Xie L, Wisse LEM, Pluta J, de Flores R, Piskin V, Manjón JV, Wang H, Das SR, Ding S-L, Wolk DA, Yushkevich PA, for the Alzheimer's Disease Neuroimaging Initiative (2019) Automated segmentation of medial temporal lobe subregions on in vivo T1-weighted MRI in early stages of Alzheimer's disease. *Hum Brain Mapp* 40, 3431–3451. [PubMed: 31034738]
- [57]. Das SR, Avants BB, Grossman M, Gee JC (2009) Registration based cortical thickness measurement. *NeuroImage* 45, 867–879. [PubMed: 19150502]
- [58]. Klooster NB, Tranel D, Duff MC (2020) The hippocampus and semantic memory over time. *Brain and Language* 201, 104711. [PubMed: 31739112]
- [59]. Collie A, Maruff P (2000) The neuropsychology of preclinical Alzheimer's disease and mild cognitive impairment. *Neuroscience & Biobehavioral Reviews* 24, 365–374. [PubMed: 10781696]
- [60]. Wagner M, Wolf S, Reischies FM, Daerr M, Wolfsgruber S, Jessen F, Popp J, Maier W, Hull M, Frolich L, Hampel H, Perneczky R, Peters O, Jahn H, Luckhaus C, Gertz H-J, Schroder J, Pantel J, Lewczuk P, Kornhuber J, Wiltfang J (2012) Biomarker validation of a cued recall memory deficit in prodromal Alzheimer disease. *Neurology* 78, 379–386. [PubMed: 22238414]

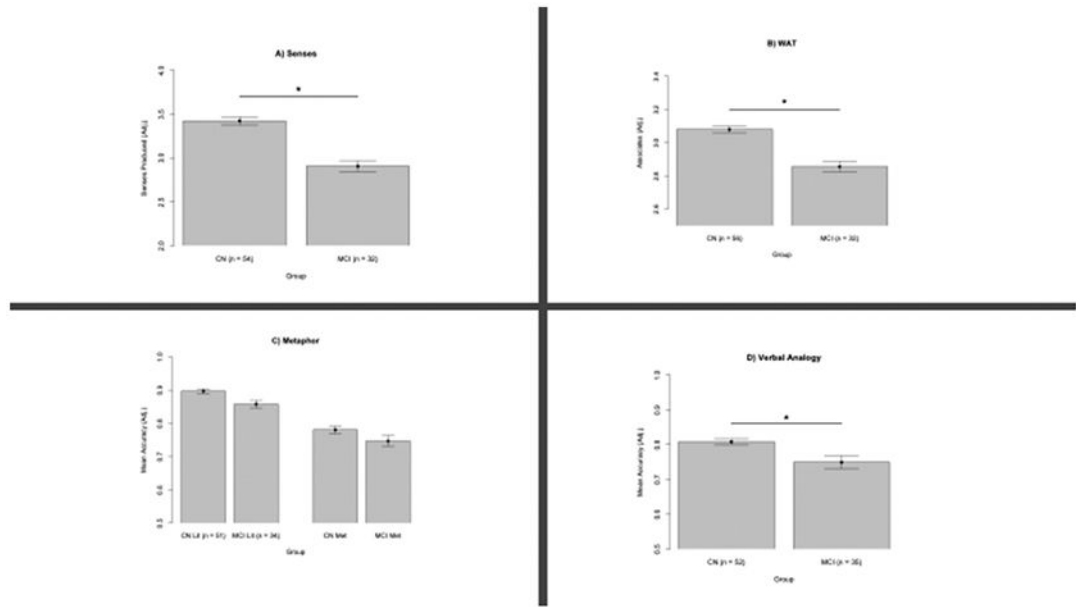
- [61]. Dubois B, Picard G, Sarazin M (2009) Early detection of Alzheimer's disease: new diagnostic criteria. *Dialogues Clin Neurosci* 11, 135–139. [PubMed: 19585949]
- [62]. Bäckman L, Small BJ, Fratiglioni L (2001) Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain* 124, 96–102. [PubMed: 11133790]
- [63]. Pike KE, Savage G, Villemagne VL, Ng S, Moss SA, Maruff P, Mathis CA, Klunk WE, Masters CL, Rowe CC (2007) -amyloid imaging and memory in non-demented individuals: evidence for preclinical Alzheimer's disease. *Brain* 130, 2837–2844. [PubMed: 17928318]
- [64]. Bäckman L, Jones S, Berger A-K, Laukka EJ, Small BJ (2005) Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology* 19, 520–531. [PubMed: 16060827]
- [65]. Grossman M, D'Esposito M, Hughes E, Onishi K, Biassou N, White-Devine T, Robinson KM (1996) Language comprehension profiles in Alzheimer's disease, multi-infarct dementia, and frontotemporal degeneration. *Neurology* 47, 183–189. [PubMed: 8710075]
- [66]. Taler V, Phillips NA (2008) Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology* 30, 501–556. [PubMed: 18569251]
- [67]. Price BH, Gurvit H, Weintraub S, Geula C, Leimkuhler E, Mesulam M (1993) Neuropsychological Patterns and Language Deficits in 20 Consecutive Cases of Autopsy-Confirmed Alzheimer's Disease. *Archives of Neurology* 50, 931–937. [PubMed: 8363447]
- [68]. Wolk DA, Das SR, Mueller SG, Weiner MW, Yushkevich PA (2017) Medial temporal lobe subregional morphometry using high resolution MRI in Alzheimer's disease. *Neurobiology of Aging* 49, 204–213. [PubMed: 27836336]
- [69]. Cardillo ER, Watson CE, Schmidt GL, Kranjec A, Chatterjee A (2012) From novel to familiar: Tuning the brain for metaphors. *NeuroImage* 59, 3212–3221. [PubMed: 22155328]
- [70]. Brier MR, Gordon B, Friedrichsen K, McCarthy J, Stern A, Christensen J, Owen C, Aldea P, Su Y, Hassenstab J, Cairns NJ, Holtzman DM, Fagan AM, Morris JC, Benzinger TLS, Ances BM (2016) Tau and A $\beta$  imaging, CSF measures, and cognition in Alzheimer's disease. *Sci Transl Med* 8, 338ra66–338ra66.
- [71]. Pontecorvo MJ, Devous MD, Navitsky M, Lu M, Salloway S, Schaerf FW, Jennings D, Arora AK, McGeehan A, Lim NC, Xiong H, Joshi AD, Siderowf A, Mintun MA, for the 18F-AV-1451-A05 investigators (2017) Relationships between flortaucipir PET tau binding and amyloid burden, clinical diagnosis, age and cognition. *Brain* 140, 748 – 763. [PubMed: 28077397]





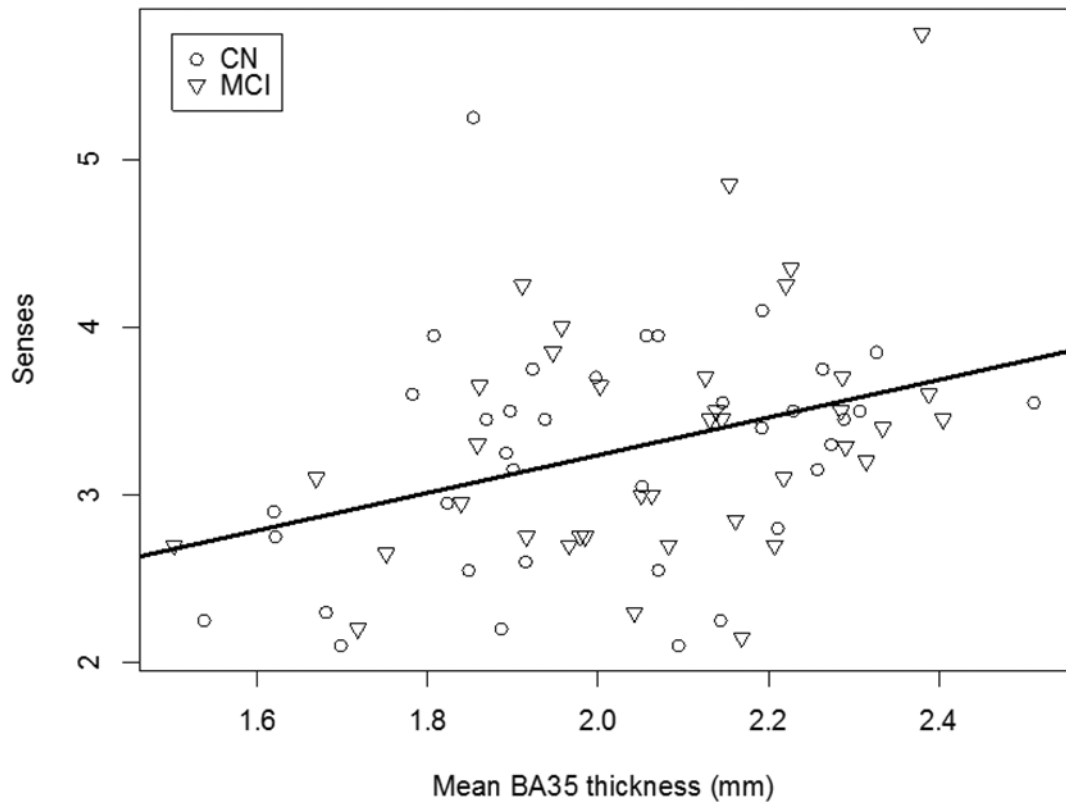
**Figure 1.**

Example stimuli from shape and color trials of visuospatial analogical reasoning task (VSA) and within-domain and cross-domain verbal analogical reasoning task (VGA), derived from Watson & Chatterjee (2012) and Green et al (2010), respectively.

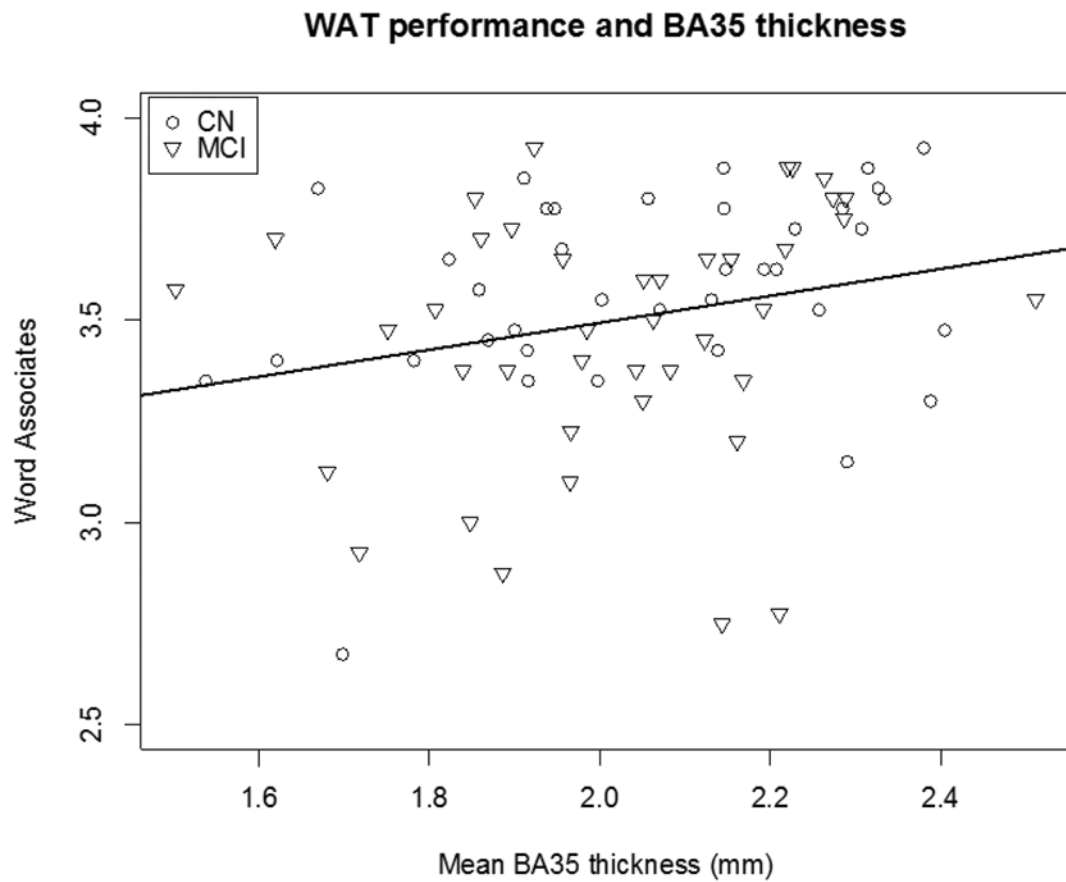


**Figure 2.** Task performance (adjusted for differences in age and level of education) from participants with normal cognition (CN) and participants with Mild Cognitive Impairment (MCI). \* $p < 0.05$ . Error bars represent standard error of the mean.

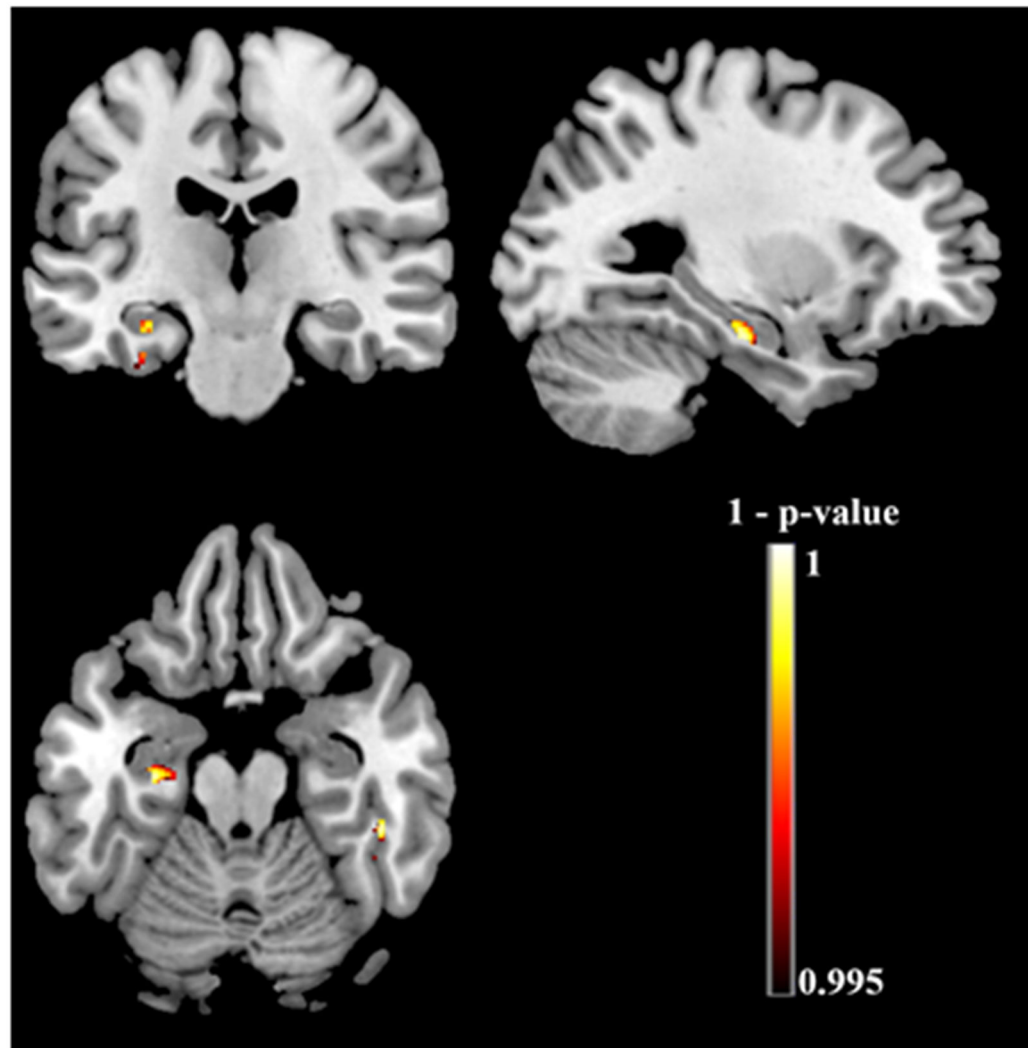
### Senses Task performance and BA35 thickness



**Figure 3.** Senses produced as a function of BA35 cortical thickness by cognitively normal participants (CN) and patients with mild cognitive impairment (MCI;  $\beta = 0.77$ ,  $p < 0.05$ ).



**Figure 4.** Mean word associates correctly identified as a function of BA35 cortical thickness by healthy comparison participants (CN) and patients with mild cognitive impairment (MCI;  $\beta = 0.191$ ,  $p < 0.05$ ) on the Word Associates Test (WAT).



**Figure 5.** Exploratory full-brain voxel-based morphometry showing relationships between performance on the Word Associates Test and brain atrophy. Clusters that meet a height threshold of  $p < 0.005$  uncorrected with threshold-free cluster enhancement and a minimum of 25 adjacent voxels are depicted.

**Table 1**

## Neuropsychological Assessment

Group	MOCA	ANIMAL Fluency	VEGETABLE Fluency	TRAILA	TRAILB	CRAFTVRS	CRAFTDVR	UDSBENTC
MCI	21.96 (3.31)	17.28 (4.87)	10.08 (3.90)	37.36 (13.21)	163.16 (188.67)	11.8 (6.16)	5.72 (6.00)	15.08 (1.68)
CN	27.10 (2.25)***	23.04 (5.39)***	14.83 (4.42)***	30.12 (7.81)**	68.13 (26.99)***	22.81 (5.32)***	20.44 (5.98)***	15.71 (1.46)

Group	UDSBENTD	DIGFORCT	DIGBACCT	MINT	UDSVERFC	UDSVERLC	Word List Memory	Word List Recall	Clock Draw
MCI	5.32 (4.59)	7.4 (2.38)	5.72 (1.86)	27.4 (3.51)	11.64 (4.53)	12.44 (5.21)	16.4 (3.84)	4.60 (2.25)	2.16 (1.55)
CN	11.85 (2.68)***	8.58 (2.29)*	7.73 (2.25)***	30.19 (2.21)***	16.21 (4.45)***	14.54 (4.52)**	24.46 (3.39)***	8.50 (1.60)***	1.21 (0.78)***

Key: MCI: Patients with Mild Cognitive Impairment. CN: Healthy participants with normal cognition. MoCA: Montreal Cognitive Assessment. CRAFTVRS: Craft Story immediate recall. CRAFTDVR: Craft story delayed recall. UDSBENTC: Benson Complex Figure Copy. UDSBENTD: Benson Complex Figure Delayed. DIGFORCT: Number span forward. DIGBACCT: Number span backward. MINT: Multilingual Naming Test. UDSVERFC: Verbal Fluency Phonemic Test 'F'. UDSVERLC: Verbal Fluency Phonemic Test 'L'.

\*  $p < 0.05$

\*\*  $p < 0.01$

\*\*\*  $p < 0.005$ .

**Table 2**

Metaphor task examples

Sentence	Target	Foil 1	Foil 2	Foil 3
<b>The uncle groaned in the other room</b>	physical suffering	generous parent	comfortable rest	broken mirror
<b>The inn groaned at the new guests</b>	crowded accommodations	audible grumble	plentiful vacancies	winding road

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3:**

## Senses Task Fixed Effects

Parameter	Estimate	Std. Error	z-value
Intercept	-2.7922	1.4251	-1.959
Senses	1.0946	0.4478	2.445*

\*  
p<0.05

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 4:**

## WAT Fixed Effects

Parameter	Estimate	Std. Error	z-value
Intercept	-8.7445	3.2699	2.674 **
WAT	2.7166	0.9355	2.904 **

\*\*  
p < 0.01

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 5:**

## Metaphor Task Fixed Effects

Parameter	Estimate	Std. Error	z-value
Intercept	-9.579	3.939	-2.432 <sup>*</sup>
Literal Trials	10.968	4.187	2.619 <sup>**</sup>

<sup>\*</sup>  
p < 0.05

<sup>\*\*</sup>  
p < 0.01.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 6:**

## Verbal Analogy Fixed Effect

Parameter	Estimate	Std. Error	z-value
Intercept	-5.226	2.825	-1.850
Task performance	6.563	3.192	2.056*

\*  
p<0.05

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 7**

## Coefficients of Linear Discriminants

<b>Test</b>	<b>Coefficients of linear discriminants</b>
<b>Literal Trials</b>	4.90143564
<b>Metaphor Trials</b>	2.22805849
<b>Verbal Analogy</b>	1.43587324
<b>Senses Test</b>	0.41808097
<b>Word Associates Test</b>	0.3592963
<b>Clock Draw</b>	0.27567142
<b>MINT</b>	0.18601609
<b>CRAFTDVR</b>	0.16658615
<b>Word List Immediate Memory</b>	0.15247097
<b>Complex Figure Copy</b>	0.10362485
<b>“F” fluency</b>	0.0914816
<b>Vegetable Fluency</b>	0.07700773
<b>CRAFTVRS</b>	0.07401966
<b>“L” Fluency</b>	0.06845429
<b>DIGFORCT</b>	0.05365811
<b>DIGBACCT</b>	0.05350568
<b>Complex Figure Delayed</b>	0.04237303
<b>MOCA</b>	0.0239047
<b>Word List Recall</b>	0.0195345
<b>TRAIL A</b>	0.01682269
<b>Animal Fluency</b>	0.0159856
<b>TRAIL B</b>	0.00088722

Key: MINT: Multilingual Naming Test. CRAFTDVR: Craft story delayed recall. CRAFTVRS: Craft Story immediate recall. DIGFORCT: Number span forward. DIGBACCT: Number span backward. UDSBENTD: Benson Complex Figure Delayed. MoCA: Montreal Cognitive Assessment.

**Table 8:**

## Senses x BA35 Fixed Effects

Parameter	Estimate	Std. Error	t-value
<b>Intercept</b>	1.20104	1.09642	1.095
<b>Group (CN)</b>	0.46299	0.15800	3.818 **
<b>Age</b>	-0.01739	0.01025	-1.697
<b>Education</b>	0.08818	0.02331	3.783 **
<b>Mean BA35</b>	0.77254	0.32650	3.238 *

\*  
p < 0.05

\*\*  
p < 0.01.

CN: Cognitively normal participants. BA35: Brodmann Area 35.

**Table 9:**

## WAT x BA35 Fixed Effects

Parameter	Estimate	Std. Error	t-value
<b>Intercept</b>	2.805618	0.493922	5.680***
<b>Group (CN)</b>	0.285037	0.071716	3.975***
<b>Age</b>	-0.005537	0.004599	-1.204
<b>Education</b>	0.036152	0.010738	3.367**
<b>Mean BA35</b>	0.328675	0.157183	2.091*

\*  
p < 0.05

\*\*  
p < 0.01

\*\*\*  
p < 0.001.

CN: Cognitively normal participants. BA35: Brodmann Area 35.

**Table 10**

Results of full-brain voxel-based morphometry analyses

cluster size	Peak p-value	x	y	z	Location
198	< 0.001	-29	-18	-20	Left hippocampus
416	< 0.001	46	-35	-21	Right inferior temporal gyrus (posterior division)
82	0.002	-32	-21	-27	Left anterior parahippocampus gyrus

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript