The Semantic Object Retrieval Test (SORT) in Amnestic Mild Cognitive Impairment

Michael A. Kraut, MD, * Barbara Cherry, PhD, † Jeffery A. Pitcock, BA, ‡ Raksha Anand, MA, § Juan Li, PhD, ‡ Lindsey Vestal, MA, ¶ Victor W. Henderson, MD, # and John Hart, Jr, MD, ‡§ ¶

Background: Between 10% and 15% of patients with the amnestic variety of Mild Cognitive Impairment (MCI) convert to Alzheimer disease (AD) per year.

Objective: Characterize cognitive markers that may herald conversion from MCI to AD and directly assess semantic memory in patients meeting criteria for amnestic MCI.

Design: Thirty-five amnestic MCI patients and 121 healthy aging controls enrolled at an Alzheimer Disease Center received a battery of standard neuropsychologic tests, and the Semantic Object Retrieval Test (SORT), a test that we have developed for the assessment of semantic memory and subsequent name production, and that has been shown to be able to differentiate between normals and patients with AD.

Results: On the basis of normative data from the SORT, the MCI subjects could be divided into 2 groups: 10 patients (29%) with a significant semantic impairment (SI+) and 25 without a semantic memory deficit (SI−). There was a significant correlation between all SORT variables and performance on the Boston Naming Test. In this MCI population, significantly impaired SORT performance was associated with a relative decrease in performance on tests of frontal lobe functions, although disruption of thalamic-related processes cannot be excluded as an etiology for semantic memory impairment.

Conclusions: The SORT is a specific test of semantic memory, and is a sensitive measure of semantic memory deficits in patients who otherwise meet criteria for amnestic MCI. Using this specific assessment tool, a significant number of MCI patients were found to have semantic memory deficits. As these patients may be early in the course of possible progression toward dementia the SORT or other tests of semantic memory, may provide important diagnostic or prognostic information in patients with MCI.

Key Words: mild cognitive impairment, semantic, memory, Alzheimer disease, dementia

(Cog Behav Neurol 2007;00:000–000)
and speech functions. This test has recently been demonstrated to be a sensitive probe of decline of semantic memory in patients with mild to moderate AD, successfully differentiating those patients from normal aging controls. The test was more sensitive than were the typically used category fluency tests, suggesting that the task of retrieving a specific object, given the constraints of 2 of the object’s component features, puts greater demands on the semantic memory system than does the less constrained task of recalling objects that belong to certain target categories in general. Evidence from functional magnetic resonance imaging (fMRI) and electrophysiologic studies suggest that performance of the SORT selectively engages thalamic structures that modulate synchronizing \( \gamma \) electrophysiologic rhythms during object recall, and medial Brodmann area (BA) 6 in the frontal lobes, which has been imputed to mediate semantic or general search mechanisms or generates an object framework from featural input.\(^6,12\) Thus, impairment of semantic object recall could result from dysfunction of thalamic-modulated synchronization of cortical regions that encode components of an object in semantic memory,\(^5,9\) from frontal dysfunction, and/or from disruption of the cortical or limbic regions encoding for the components of the object.\(^7\) The following describes our findings upon administration of the SORT to a cohort of amnestic MCI patients, the performance of the SORT in comparison with other cognitive tests commonly used in the evaluation of patients with suspected dementia, and the implications for diagnostic and prognostic issues in amnestic MCI.

**METHODS**

**Subjects**

We studied 35 consecutive patients who enrolled in the Alzheimer’s Disease Center (ADC) at UAMS between September, 2001 and October, 2004, and who were classified for research purposes as having amnestic MCI.\(^1\) These criteria, as adapted for use at the ADC and adjudicated at a diagnostic consensus conference that included behavioral neurologists, psychologist, and other health care professionals were the following: (1) subjective memory complaint or a family member’s history of forgetfulness, (2) no more than minimal impairment in activities of daily living, (3) normal general cognitive function (on the basis of history and neuropsychologic test results including tasks of working memory, executive function, visuospatial skills, naming, etc), (4) abnormal memory for age on the basis of neuropsychologic test results (WMS-III Logical Memory and word list tasks), and (5) no dementia on the basis of DSM-IV criteria. In addition, patients had a Clinical Dementia Rating score of 0.5-1. SORT results were not used in determining whether patients met MCI criteria.

Patients were between 49 to 88 years of age, in good health, with no evidence of focal neurologic signs and a recent brain imaging study, demonstrating no intracranial pathology likely to impair cognition.

This study was conducted according to the Good Clinical Practice Guidelines, The declaration of Helsinki, and the US Code of Federal Regulations. Written and informed consent was obtained from all participants and their caregivers according to the rules of the HRAC of UAMS.

**Neuropsychologic Tests**

The following tests were administered as part of the neuropsychologic battery:

1. (1) Mini-Mental State Examination.
2. (2) North American Adult Reading Test.
4. (4) Symbol Digit Modalities Test.
5. (5) Category Fluency Test (animals).
6. (6) Controlled Oral Word Association Test (FAS).
7. (7) Word List Learning (Wechsler Memory Scale III).
8. (8) Trail Making Tests A and B.
10. (10) Praxis Battery from the Boston Diagnostic Aphasia Examination.
11. (11) Multilingual Aphasia Examination Token Test—Modified.
12. (12) Word List Learning (Wechsler Memory Scale III).
14. (14) Category Fluency Test (animals).
15. (15) Praxis Battery from the Boston Diagnostic Aphasia Examination.
16. (16) Trail Making Tests A and B.
17. (17) Stroop Interference Test.
18. (18) Semantic Memory (name production).
20. (20) Geriatric Depression Scale (GDS). Semantic Memory (and name production).
21. (21) SORT.\(^4\) The stimuli consist of verbally presented word pairs. The words for this task are all features of objects. There are 2 types of word pairs: (a) 16 pairs, where the 2 words describing features of an object combine to elicit an object that was not presented (eg, the words desert and humps, which produce the object camel), and (b) 16 word pairs that do not combine to recall an object not presented and are semantically unrelated (eg, humps and “alarm”). The same feature words used in the object recall pairs comprise the stimuli in the unrelated pairs, but are repaired with a semantically unrelated word (eg, humps and alarm).

The participants in the study were instructed immediately before testing as to the meaning of “the 2 words combine together to make you think of a particular object.” The participants were instructed to say or signal yes or no if the words combine together to recall an object. They were further instructed that for the word pairs resulting in object recall, to later provide the name of the object.

Scoring for the SORT reflects correct answers, and also 2 types of semantic memory errors on the object
recall aspect of the test—false positives (or “overbinds”) and false negatives (“underbinds”).

The name production aspect of the task occurs after the semantic memory component and was scored for errors of “substitutions” which is when correct recall occurs, but the name provided by the subject is not correct and “omissions” when the subject reports that they recall an object but do not know its name. Analyzed scores consist of total number of correct names and number of errors for each subtype (substitutions and omissions).

The participants also received with these instructions a set of standardized practice items to ensure that subject understood the task. If it was evident from their responses in the practice test that the subject did not understand the nature of the task or comprehend the instructions, the test was not administered (no instances in this MCI population). If it was considered from the practice sessions that the participant understood the task, the entire test was administered.

**Procedures**

All testing was performed by trained neuropsychologic testing technicians over 1 or more testing sessions. Administration of the SORT took approximately 10 minutes. The technician recorded all responses the subject made and scored the tests after the testing sessions.

**RESULTS**

(1) The healthy aging control subjects in the ADC also received the same set of neuropsychologic tests, including the SORT. The normative data from this population are reported elsewhere. In brief, 121 subjects classified as healthy aging control subjects had a mean age of 72.6 (standard deviation (SD) = 6.0), education of 15.3 years (SD = 2.5), and a mean SORT object recall score of 29.4 (SD = 2.1), with 2.0 (SD = 2.0) false-positive and 0.6 (SD = 1.0) false-negative errors. Using a conservative estimation of a semantic object recall impairment as being a test score of greater than or equal to 2 SD from the mean (cut-off of 25 or less correct), 10 of the 35 MCI subjects demonstrated significant deficits in semantic object recall compared with healthy aging controls.

These 2 groups, MCI patients with a semantic memory impairment (SI+) and those without an impairment (SI−), did not differ significantly on age, education, Mini-Mental State Examination score, FSIQ as estimated by the North American Adult Reading Test, GDS score (Table 1), or possession of an apolipoprotein E 4 allele.

Initial analyses of variance (ANOVA) were conducted with all 3 groups (SI+, SI−, controls) as levels of the independent variable and SORT variables as dependent variables to assure that subsequent t tests were appropriate. These were conducted as both ANOVAs and analyses of covariance with education, FSIQ, and mood (GDS) as covariates (age did not differ among the 3 groups). Both sets of analyses demonstrated the same pattern of significant differences for all measures of the SORT and are reported here as ANOVAs: memory total

<table>
<thead>
<tr>
<th>Variable</th>
<th>MCI SI+ (n = 10)</th>
<th>MCI SI− (n = 25)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>74.1 (6.9)</td>
<td>73.0 (8.4)</td>
<td>ns</td>
</tr>
<tr>
<td>Education (y)</td>
<td>13.2 (3.8)</td>
<td>14.6 (2.4)</td>
<td>ns</td>
</tr>
<tr>
<td>MMSE (Serial 7’s)</td>
<td>25.0 (4.2)</td>
<td>26.4 (2.5)</td>
<td>ns</td>
</tr>
<tr>
<td>MMSE (Spelling)</td>
<td>26.3 (3.2)</td>
<td>26.5 (2.4)</td>
<td>ns</td>
</tr>
<tr>
<td>FSIQ (NAART)</td>
<td>93.1 (27.4)</td>
<td>107.1 (8.6)</td>
<td>ns</td>
</tr>
<tr>
<td>GDS</td>
<td>5.7 (4.1)</td>
<td>5.8 (5.1)</td>
<td>ns</td>
</tr>
</tbody>
</table>

*p = 0.14 when adjusting for unequal variances.

**TABLE 2.** Means and SD for SORT Variables for MCI Impaired and Unimpaired MCI Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>MCI SI+</th>
<th>MCI SI−</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory Total (max.32)</td>
<td>22.0 (2.7)</td>
<td>28.9 (1.5)</td>
<td>29.4 (2.1)</td>
</tr>
<tr>
<td>False Positive (max.16)</td>
<td>8.1 (4.2)</td>
<td>2.2 (1.3)</td>
<td>2.1 (2.0)</td>
</tr>
<tr>
<td>False Negative (max.16)</td>
<td>1.9 (2.0)</td>
<td>0.9 (0.9)</td>
<td>0.6 (1.0)</td>
</tr>
<tr>
<td>Correct Names (max.16)</td>
<td>11.1 (2.8)</td>
<td>14.4 (1.4)</td>
<td>14.8 (1.2)</td>
</tr>
<tr>
<td>Substitutions (max.16)</td>
<td>3.0 (2.9)</td>
<td>0.8 (0.8)</td>
<td>0.6 (0.8)</td>
</tr>
</tbody>
</table>

*All SORT variables significantly different for MCI SORT impaired (SI+) vs. unimpaired (SI−).

**TABLE 3.** Means and SD for Demographic Variables for SORT Impaired and Unimpaired MCI Participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>MCI SI+</th>
<th>MCI SI−</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>74.1 (6.9)</td>
<td>73.0 (8.4)</td>
<td>ns</td>
</tr>
<tr>
<td>Education (y)</td>
<td>13.2 (3.8)</td>
<td>14.6 (2.4)</td>
<td>ns</td>
</tr>
<tr>
<td>MMSE (Serial 7’s)</td>
<td>25.0 (4.2)</td>
<td>26.4 (2.5)</td>
<td>ns</td>
</tr>
<tr>
<td>MMSE (Spelling)</td>
<td>26.3 (3.2)</td>
<td>26.5 (2.4)</td>
<td>ns</td>
</tr>
<tr>
<td>FSIQ (NAART)</td>
<td>93.1 (27.4)</td>
<td>107.1 (8.6)</td>
<td>ns</td>
</tr>
<tr>
<td>GDS</td>
<td>5.7 (4.1)</td>
<td>5.8 (5.1)</td>
<td>ns</td>
</tr>
</tbody>
</table>

*p = 0.14 when adjusting for unequal variances.

MMSE indicates Mini-Mental State Examination; NAART, North American Adult Reading Test.

**TABLE 1.** Means and SD for Demographic Variables for SORT Impaired and Unimpaired MCI Participants

It is important to note that those with a significant semantic memory impairment did not seem to have any difference in performance on tests of auditory compre-
hension (Token Test, \( t \) test, \( P > 0.70 \)), supporting the claim that a comprehension deficit does not account for differences in SORT performance.

(3) Naming error types that were present to a significant degree were substitutions. The other score that reflected naming performance in the normal aging study was the total number of correct names produced [termed Correct Names (n = 16)]. The Correct Names score in the healthy aging control group was 14.8 average correct (SD = 1.2). The SI+ group made significantly more substitution errors (mean 3 errors, SD = 2.9) than did the SI− (mean 0.8, SD = 0.8) and the healthy aging control groups (substitution error mean of 0.6, SD = 0.8) (\( t \) tests, \( P < 0.001 \) (Table 2). The SI+ group also produced significantly fewer correct names (mean 11.1 correct, SD = 2.8) compared with the SI− (mean 14.4 correct, SD = 1.4) and the healthy aging group (mean 14.8 correct, SD = 1.2) (\( t \) tests, \( P < 0.005 \)). Other test results were similar.

(4) Group performances were compared on other tasks in the neuropsychologic battery. There were significant differences between the SI+ and SI− groups on a subset of these tasks: BNT, Repetition of Low Probability Phrases, Stroop Interference Test, Matrix Reasoning, Necker Cube Figure Copying, Word List Learning 1 (trial 4 only), and for Logical Memory, immediate recall of Story B only (\( t \) statistic comparisons between SI+ and SI− significant, \( Ps < 0.05 \); 3-group ANOVAs with controls, SI+, and SI− significant, \( Ps < 0.001 \)). Other test results were similar.

(5) We examined the association between all of the SORT variables (including both memory and name production measures) and the neuropsychologic testing results. There was a significant correlation between all of the SORT variables and performance on the BNT, a test which engages some aspects of lexical-semantics and several other cognitive functions. Scores for Memory Total and Correct Names were positively correlated with BNT (\( rs = 0.66, Ps < 0.001 \)), whereas False Positive, False Negative, and Substitution scores were negatively associated with BNT (\( rs = 0.36 \) to \( -0.56, Ps < 0.05 \)). Using the BNT score as a covariate and MCI group (SI+ and SI−) as the independent variable, the dependent variables that are still significantly differentiate between SI groups are semantic memory total score [\( F(2,30) = 48.53, P < 0.001 \)], false positives [\( F(2,30) = 18.82, P < 0.001 \)], correct names [\( F(2,30) = 16.25, P < 0.001 \)], and substitutions [\( F(2,30) = 8.70, P < 0.001 \)] (dependent variable that does not demonstrate significant differences between the SI+ and the SI− groups is memory false negatives [\( F(2,30) = 3.01, P > 0.05 \)]. These findings indicate that SORT assesses cognitive functions that differ in part from those assessed with the BNT.

**DISCUSSION**

Typically, AD presents initially with impairment in episodic memory, particularly delayed recall of newly learned verbal material. This finding is often among the first symptoms of AD, although frontal lobe/executive dysfunction is also commonly seen early in the disease course.\(^{13}\) However, this is not always the case, and the anatomic localization of the frontal lobe structures involved has not been delineated.

The present study describes a group of individuals with a diagnosis of amnestic MCI who also have a significant semantic memory deficit (greater than 2 SD from the mean of age-matched healthy controls on the SORT task). We used conservative criteria for SORT impairment on the basis of norms derived from age-appropriate, healthy controls, and thus it is unlikely that the semantic memory deficits were present premorbidly. Our findings do not show whether semantic memory deficits occurred contemporaneously or after the onset of episodic memory deficits, but results indicate that semantic deficits are a common early feature of patients who otherwise meet criteria for amnestic MCI.

The anatomic regions identified by functional imaging and electrophysiology studies as associated with the semantic retrieval process are medial BA6 in the frontal lobes, the thalamus (including the pulvinar and dorso medial nuclei), and the basal tempo ro-occipital visual memory system.\(^{4-8}\) Given the imputed role of BA6 in refining semantic search criteria or in selective engagement of appropriate target objects and their components (eg, features, category),\(^{6,13}\) dysfunction could possibly lead to false-positive errors in object recall. The thalamus is postulated to unite or bind features and other object components together via \( \gamma \) rhythm synchronization of the cortical regions encoding these aspects of an object.\(^{6,8}\) Thalamic dysfunction in a degenerative state, which would most likely result in decreased recall overall and associated false-negative errors. However, an “over-binding” of features resulting in the false-positive object recall, which is what was detected in the subset of SI+ MCI patients, is more consistent with frontal lobe dysfunction, likely involving medial BA6. This is supported by the neuropsychologic performance differences between the SI+ and SI− MCI groups. The SI− group performed significantly less well on 7 tests, including Matrix Reasoning, the BNT, Repetition of Low Probability Phrases, and the Stroop Interference Test. The Matrix Reasoning subtest of the WAIS-III has associations with abstraction and fluency that are frontal lobe-related functions.\(^{16}\) The BNT has been employed to engage brain regions including the superior medial frontal lobes at BA6,\(^{7,17}\) as evidenced by signal changes detected during fMRI. Conversely, impaired performance on this test has been associated with atrophy detected by voxel-based morphometric measures in patients with frontotemporal dementia. FMRI-based studies of subjects performing a phrase repetition task have also exhibited brain activation in the medial BA6 region.\(^{18}\) Performance on the Stroop Interference Test has consistently been associated with activation in the rostral anterior cingulate gyrus neighboring medial BA6\(^{19}\) as revealed with fMRI, while impaired performance has been linked to bilateral lesions of the superior medial frontal lobes.\(^{20}\) Similar to the above-specified tests, the SORT elicits fMRI-detected
evidence for activation in medial BA6, implicating dysfunction of this common region as a major factor in poor performance on the SORT.

Focal cerebrovascular lesions in the thalamus have resulted in impaired SORT performance.\(^9\) Delayed or decreased thalamic modulation of synchronizing \(\gamma\) rhythms associated with semantic memory would likely lead to a decrease in object recall and thus more false-negative responses. However, thalamic “dysmodulation” that improperly gates \(\gamma\) rhythm synchronization and thus allows for synchronization to occur even when the features are unrelated, could potentially result in false-positive object recall. Numerous studies demonstrating thalamic pathology in AD,\(^{21}\) including the pulvinar\(^{22}\) and dorsomedial nucleus,\(^{23}\) and evidence that AD and MCI patients in general exhibit less overall baseline \(\gamma\) frequency-band electro encephalogram activity than do controls,\(^{24}\) suggests that thalamic dysfunction is also a possible etiology for the present findings. The plausibility that the semantic memory deficit could be associated with thalamic pathology in MCI is supported by the findings that thalamic pathologic changes can be detected relatively early in the course of AD\(^{22}\) and that development of verbal learning and memory deficits in AD have been associated with the occurrence of both medial temporal lobe and thalamic damage.\(^{25}\)

Regardless of the pathophysiologic mechanisms that result in decrements in object recall, the SORT proved a useful measure of semantic memory deficits in MCI. Other tasks typically used in the evaluation of dementia assess some aspects of semantic processing, but also engage a variety of other cognitive processes unrelated to semantics. It is clear from this and other studies that semantic memory is an essential cognitive domain to assess in degenerative disease, particularly because it has been shown to be impaired early in these conditions. Thus, assessments of the semantic memory domain need to be sensitive and specific in this elderly population. The SORT is a specific measure of semantic memory functions, with few ancillary nonsemantic cognitive components necessarily engaged by the task.

A major distinction between the MCI SI+ patients, MCI SI− patients, and healthy aging controls is the relatively large number of false-positive memory errors and production of a name for each of these falsely remembered objects. These names at times were related to objects that were nonexistent (eg, “mane” and “wings” resulted in “lions”). These findings represent an unusual “positive” symptom for neurologic disease states in general, particularly a degenerative state. Atypical frontal lobe behavioral abnormalities and hallucinations can be construed as “positive” symptoms and appear in degenerative diseases such as AD, but typically later in the course of the disease, or in frontal dementia. There are examples of neuropsychiatric symptoms in MCI,\(^{26}\) but these are typically “negative” symptoms (eg, mood disturbance, apathy, etc). These neuropsychiatric symptoms have been accounted for in a variety of ways, with the most prevalent proposal being frontal dysfunction resulting in failure to inhibit or monitor behavioral impulses. An analogous account can be provided for the false-positive semantic memories in these patients, with BA6 frontal lobe dysfunction leading to impaired monitoring or inhibition of the numerous choices that are considered in semantic object recall, leading to not rejecting incorrect object choices.

Studies have reported impairments in verbal episodic memory in amnestic MCI patients on tasks such as delayed word list recall, delayed story recall, and associative learning (see Ref.\(^{27}\) for review).\(^{28}-^{30}\) Recent investigations have also suggested that semantic memory impairments are present in subsets of individuals with amnestic MCI.\(^{31-32}\) These semantic memory declines have been posited on the basis of performance in the category fluency task\(^{36}\) (ability to generate examples within a specific semantic category), which in most cases was found to be inferior compared with the performance demonstrated by normal controls but better than that of individuals with clinical AD.\(^{31,33-35}\) Although the category fluency task clearly requires retrieval and selection of exemplars from a group of items organized as a lexical-semantic category, there are additional cognitive operations that are engaged by this task that are not specific semantic operations. Thus, impairments in this task may not necessarily represent semantic deficits unless confirmed on other tasks. For example, Adlam et al\(^{37}\) reported that a group of amnestic MCI performed significantly differently than a normal control group on the category fluency task and tests of object knowledge (matching to recipient, function, and action), providing further evidence of a semantic deficit on category fluency. Other tests of semantic memory have not demonstrated sensitivity in detecting early semantic deficits in dementia or amnestic MCI, or at least do not provide strong supporting evidence of a semantic deficit. The Pyramids and Palm Trees tests, which has been shown to be specific to impairments in associative semantic knowledge, was not found to be as sensitive to the semantic processes that are disrupted early in amnestic MCI as the category fluency test.\(^{33}\)

It seems that the types of semantic processes that are differentially impaired in amnestic MCI are those where the target of a semantic search is a specific object, more so than a generalized categorical search. Duong et al\(^{22}\) demonstrated in a variety of semantic memory tasks, finding that those tasks requiring accessing information about a specific object were impaired in a group of amnestic MCI patients compared with a group of normal controls.

The SORT task as we have implemented it can be applied to individual patients, with a clinically derived cut-off for impairment on the basis of normal aging norms. In addition, the SORT probes the process of semantic search for a specific object from its component features, which appears to be a type of semantic search (for a specific object) shown to be disrupted in some patients with amnestic MCI.
Impaired performance on the SORT has not correlated thus far with other markers of cognitive progression noted in the previous studies of MCI progression to AD, including the presence of APoE 4 alleles in this cohort. This likely reflects the many pathophysiologic processes that influence the progression from MCI, or perhaps the size of the cohort we evaluated. Further exploration of early markers of cognitive decline, either behaviorally, anatomically, genetically, or in combination, may provide additional insights into early detection of degenerative decline and treatments in MCI.

ACKNOWLEDGMENTS

The authors thank Dr D. Tib for helpful suggestions and Kathy Tyler-King, RN, Andrew Howard, Dr Roger Williams for invaluable assistance in data collection.

REFERENCES