Discriminating Age-Associated Memory Impairment From Alzheimer’s Disease

James R. Youngjohn, Glenn J. Larrabee, and Thomas H. Crook III
Memory Assessment Clinics, Inc., Scottsdale, Arizona

The authors attempt to provide a better understanding of the differences between the normal memory decline characteristic of age-associated memory impairment (AAMI) and the pathological decline typical of mild Alzheimer’s disease (AD). Batteries of traditional memory tests and computer-simulated everyday-memory tests discriminated between the 2 groups, which were matched on age, gender, and education, with reasonable degrees of accuracy (87.5% and 88.4%, respectively). False positives were the most frequent classification errors when using either battery. These results indicate that it is possible to use ecologically valid memory assessment paradigms without sacrificing discriminant validity. The clinical significance of discriminating mild AD from AAMI is discussed.

The differentiation of the pathological cognitive decline seen in individuals suffering from mild Alzheimer’s disease (AD) from the nonpathological decline associated with normal aging continues to be a major focus of investigation (e.g., Bayles, Boone, Tomoeda, Slauson, & Kaszniak, 1989; Becker, Huff, Nebes, Holland, & Boller, 1988; Morris & Fulling, 1988; Spinnler & Sala, 1988). It is generally agreed that memory decline is one of the earliest markers of AD, as well as one of the most profoundly affected functions in the mild and moderate stages of AD (Larrabee, Largen, & Levin, 1985; Masur et al., 1989; Salmon, Granholm, McCullough, Butters, & Grant, 1989). However, the fact that many cross-sectional studies have revealed a continuum in the level of memory and cognitive functioning displayed by an elderly population has complicated the task of early identification of AD (Mittenberg, Seidenberg, O'Leary, & DiGiulio, 1989; Pfeffer, Afifi, & Chance, 1987).

It is widely recognized that many cognitive functions often decline in the later decades of adulthood, particularly learning and memory (Fozard, 1985; Poon, 1985). Until recently, there has been no generally accepted diagnostic classification for persons who experience such a decline. Kral (1962, 1966) introduced the term benign senescent forgetfulness to describe otherwise healthy elderly individuals who experience cognitive declines relative to their age peers. More recent studies have validated this concept (Larrabee, Levin, & High, 1986). This nosological framework fails, however, to address the larger number of individuals experiencing memory loss associated with normal developmental processes.

A workgroup sponsored by the National Institute of Mental Health proposed diagnostic criteria for the classification of age-associated memory impairment (AAMI; Crook, Bartus, et al., 1986). This nosological category refers to persons at least 50 years of age who both complain of memory impairment in tasks of daily life and have an objective memory test performance at least 1 standard deviation below the mean established for young adults.

Attempts to discriminate AAMI from AD thus far have relied primarily on cutoff scores on mental status examinations (e.g., Mini-Mental State Examination; Folstein, Folstein, & McHugh, 1975) or dementia rating scales (e.g., Global Deterioration Scale; Reisberg, Ferris, DeLeon, & Crook, 1982) or both. Reisberg and colleagues proposed combining Folstein’s (1983) cutoff score of 23 on the Mini-Mental State Examination (MMSE) with a cutoff score of 4 on the Global Deterioration Scale (GDS) to identify individuals with pathological cognitive decline (AD). Longitudinal follow up revealed that individuals whose scores were at or below these scores had rather malignant prognoses, whereas those whose scores were above the cutoffs tended to remain relatively stable, thus adding support for the construct of AAMI (Reisberg et al., 1986).

A recent longitudinal epidemiological study (Lane & Snowdon, 1989) reported relatively high prevalence and incidence rates for AAMI (prevalence: 34.93%, SE = 4.54; incidence: 6.63%, SE = 9.41, per annum), as compared with AD (prevalence: 13.01%, SE = 7.11; incidence: 3.06%, SE = 9.79, per annum). The high incidence and prevalence rates for AAMI emphasize the need for further understanding of this area. Although the Lane and Snowdon study and the Reisberg et al. (1986) work suggest that most subjects with AAMI do not progress to AD, there has been no large-scale study directed at analyzing AD and AAMI differences.

Using psychometric instruments and a discriminant analysis procedure, some investigators have reported highly accurate classification of mild AD patients and “normal” control subjects (Storandt, Botwinick, Danziger, Berg, & Hughes, 1984). This procedure identifies those measures that best discriminate between groups, assigns weights to those measures, and then predicts group membership for individuals by using a regression equation that includes their weighted test scores. Storandt et al. (1984) used this procedure to correctly classify 98%
of their subjects. The measures they examined included tests of memory, speeded psychomotor performance, and language.

This approach has not been uniformly successful, however. Using the same traditional neuropsychological tests, Storandt and Hill (1989) attempted to extend their original findings by discriminating between mild-AD patients, very-mild-AD patients, and normal control subjects. They were unable to achieve an acceptable degree of specificity with respect to the very-mild-AD patients, although they remained highly accurate in their classification of mild-AD patients and normal control subjects.

A factor of increasing importance in the psychometric differentiation of neurological disorders in general, and Alzheimer's disease in particular, is the ecological validity of assessment procedures (Larrabee & Crook, 1988; Neisser, 1982). The lack of ecological validity in memory testing (i.e., the measurement of memory functions that are directly relevant to the individual's everyday environment) has been identified as a serious shortcoming of current psychometric paradigms for assessing the elderly (Cunningham, 1986; Erickson & Scott, 1977; Mayes, 1986; Neisser, 1982). Although numerous investigations using traditional measures have demonstrated their sensitivity to brain dysfunction, problems may arise in evaluating the elderly, who typically are encountering standardized testing for the first time in their lives. As Cunningham (1986) noted, face validity becomes very important in dealing with aged individuals because they are not familiar with laboratory procedures and may resist tasks they view as trivial or ridiculous.

Crook and colleagues have attempted to create more ecologically and face-valid memory measurement instruments by using computer technology to simulate everyday-memory tasks (Crook & Larrabee, 1988; Crook, Salama, & Gobert, 1986; Larrabee & Crook, 1989). The test-design strategy has been to combine realistic, everyday stimuli with current memory measurement paradigms. The Memory Assessment Clinics (MAC) battery thus combines improvements in face and ecological validity with the advances in standardization and ease of data storage and analysis that are possible in computerized assessment. The battery was initially devised for measuring treatment response to candidate pharmaceutical compounds in age-related cognitive decline (Larrabee & Crook, 1988), and recently these measures have demonstrated positive treatment effects of an experimental phospholipid compound on AAMI (Crook, Tinklenberg, et al., 1991).

The purpose of the present investigation was to gain a better understanding of the differences between AAMI and mild AD. Specifically, we were interested in whether the computer-simulated everyday-memory battery would be useful in an application for which it was not originally developed (i.e., the discrimination of mild AD from AAMI). We examined only the mildly demented subsample of our entire AD group. To accomplish this, we selected those individuals whose MMSE scores ranged from 20 to 23. We also compared the relative discriminative validity of the everyday-memory battery with that of a battery of traditional memory tests.

Method

Participants

These data were collected as part of a larger, double-blind, placebo-controlled, FDA-supervised study of the treatment effects of a candidate pharmaceutical compound on AD and AAMI. Only the baseline data (drug- and placebo-free) were considered in the present analysis.

The AD and AAMI groups were matched on the demographic variables of age, gender, and years of education. All subjects were matched perfectly on gender and within 2 years on both age and education. Of the 56 matched AD-AAMI pairs, 33 were perfectly matched on all three variables.

The AD group consisted of 56 volunteers (25 men and 31 women) between the ages of 50 and 75 years (mean age = 65.63, SD = 6.57; mean years of education = 13.79, SD = 3.07). All subjects in the mild-AD group met the criteria for probable (presumptive) Alzheimer's disease established by the National Institute of Neurological and Communicative Disorders and Stroke and by the Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA; McKhann et al.,

Table 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>AD (M, SD)</th>
<th>AAMI (M, SD)</th>
<th>Univariate F(1,110)</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logical Memory</td>
<td>3.08 1.49</td>
<td>7.07 2.82</td>
<td>87.91* .44</td>
<td>.46</td>
</tr>
<tr>
<td>Hard Pairs Learning</td>
<td>.86 1.67</td>
<td>4.95 3.01</td>
<td>79.03* .42</td>
<td>.32</td>
</tr>
<tr>
<td>Easy Pairs Learning</td>
<td>5.83 1.89</td>
<td>7.95 1.73</td>
<td>38.10* .26</td>
<td>.24</td>
</tr>
<tr>
<td>Benton Revised Visual Retention</td>
<td>2.41 1.39</td>
<td>5.05 1.98</td>
<td>67.15* .38</td>
<td>.48</td>
</tr>
<tr>
<td>Total correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. AD = Alzheimer's disease; AAMI = age-associated memory impairment.

* \( p < .001 \).
subtest of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) as an indicator of adequate premorbid intellectual competence; and (e) a score 1 standard deviation below the mean established for healthy elderly individuals on at least one of the following memory instruments: the Benton Revised Visual Retention Test (BVRT; Benton, 1974; 4 or fewer correct); the Logical Memory subtest of the Wechsler Memory Scale (WMS; Wechsler, 1945; 5 or fewer correct); and the Paired Associates subtest (Total) of the WMS (9 or fewer correct).

The AAMI group also consisted of 56 volunteers (25 men and 31 women) between the ages of 50 and 75 years (mean age = 65.50, SD = 6.44; mean years of education = 13.80, SD = 2.60). All participants met the inclusion and exclusion criteria for AAMI (Crook et al., 1986). In addition to being at least 50 years old, they had subjectively noticed a decline in memory relative to their younger adult years and met the following psychometric criteria: (a) at least one performance that was at least 1 standard deviation below the mean for young adults on either the BVRT, Logical Memory subtest of the WMS, or the Paired Associate Learning subtest (hard pairs) of the WMS (cutoff scores of 7 or fewer, 6 or fewer, and 6 or fewer correct, respectively); (b) a raw score of at least 32 (scaled score of 9 or higher) on the WMS Vocabulary subtest; and (c) a score of 24 or higher on the MMSE to exclude dementia. Exclusion criteria included any medical, psychiatric, or neurological conditions that could affect cognitive functioning that were revealed in the medical history, clinical examination, brain CT, and comprehensive laboratory studies.

**Apparatus**

The computer-simulated everyday-memory tests of the Memory Assessment Clinics battery are administered using a Sony 19" PVM 1910 "Personal Touch" touchscreen color monitor interfaced with an AT&T 6300 computer equipped with a 20 megaray hard-disk drive, a Pioneer LDV-6010 laser-disk player, and customized computer graphics hardware. The tester is present throughout the session and sits behind "Personal Touch" touchscreen color monitor interfaced with an AT&T LDV-6010 laser-disk player, and customized computer graphics hardware. The tester is present throughout the session and sits behind the computer.

**Procedure**

All participants in both the AD and AAMI groups were administered the following traditional memory tests in the standard format:

### Table 2

**Classification Rates of Traditional Memory Tests**

<table>
<thead>
<tr>
<th>Actual group membership</th>
<th>Predicted group membership</th>
<th>Total actual cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AD</td>
<td>AAMI</td>
</tr>
<tr>
<td>AD</td>
<td>52</td>
<td>92.9</td>
</tr>
<tr>
<td>AAMI</td>
<td>10</td>
<td>17.9</td>
</tr>
<tr>
<td>Total predicted cases</td>
<td>62</td>
<td>50</td>
</tr>
</tbody>
</table>

*Note.* AD = Alzheimer's disease; AAMI = age-associated memory impairment.

### Table 3

**Computer-Simulated, Everyday Memory Test Performance in AD and AAMI**

<table>
<thead>
<tr>
<th>Measure</th>
<th>AD</th>
<th>AAMI</th>
<th>Univariate</th>
<th>Multivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>F(1,110)</td>
</tr>
<tr>
<td>4 Name-Face</td>
<td>.64</td>
<td>.72</td>
<td>2.73</td>
<td>1.26</td>
</tr>
<tr>
<td>6 Name-Face*</td>
<td>.55</td>
<td>.74</td>
<td>2.14</td>
<td>1.48</td>
</tr>
<tr>
<td>Misplaced Objects (total found)</td>
<td>8.75</td>
<td>3.36</td>
<td>13.45</td>
<td>2.59</td>
</tr>
<tr>
<td>Recognition of Faces Delayed Non-Match-to-Sample (total correct)</td>
<td>10.11</td>
<td>5.54</td>
<td>17.23</td>
<td>4.40</td>
</tr>
</tbody>
</table>

*Note.* AD = Alzheimer's disease; AAMI = age-associated memory impairment.

* p < .01. ** p < .001.
Logical Memory and Paired Associate Learning from the WMS and the BVRT.

The computerized everyday-memory tests were administered in a standardized manner in a carefully controlled clinical setting. Descriptions of those tests considered in the present investigation follow.

Misplaced Objects test. This is a test of object location recall. Subjects are required to place 20 common objects into the schematic representation of a 12-room house (maximum of 2 objects per room), using the touchscreen. Delayed recall is tested at 40 min by requesting subjects to press the screen image of the room in which they placed each of the objects. A second attempt is allowed if they should miss on the first attempt, to closely simulate the common activity of looking for a lost article. In this investigation, we examined the total number of objects correctly located during both attempts. Previous studies have suggested that the Misplaced Objects test is a measure of verbal−visual associative memory (Crook, Youngjohn, & Larrabee, 1990). It is structurally similar to a paradigm used in animal-model pharmacological research (Bartus, Dean, & Beer, 1983; Bartus, Fleming, & Johnson, 1978; Bartus & Johnson, 1976).

Name−Face Association. In this test, subjects are presented with live, color videorecordings (stored on a laser disk) of individuals introducing themselves by common first names. After a series of introductions, recall is assessed by showing the images of the same individuals in a different order and asking the subject to provide the name of each person. To provide the subject with expressional and acoustical cues available in daily life, individuals who appear on screen during the recall phase say the name of the city in which they reside. In the present investigation, we examined immediate recall after series of four and six different introductions. Name−Face Association has been shown to be a sensitive measure of age-related memory decline (Crook & West, 1990). Recognition of Faces-Delayed Non-Matching-to-Sample. On the first trial of this test, subjects are presented with a single facial photograph on the touchscreen monitor and asked to touch the image of the face. On each of the 24 subsequent trials, a new face is added to the array, with the subject being required to identify the new face added by touching the image of it on the monitor. Each trial is separated from the preceding trial by an 8-s interval, during which the screen is black. Feedback is provided on each trial in the form of a red square that appears momentarily around the photograph if it is correctly identified. Previous factor-analytic studies have demonstrated this test to be a relatively robust measure of visual memory (Larrabee & Crook, 1989). This test is based on an earlier version, which used the non-matching-to-sample paradigm with household objects and was shown to discriminate between normal and demented elderly subjects (Flicker, Ferris, Crook, & Bartus, 1987). The delayed non-matching-to-

sample paradigm is a memory measure comparable in construct validity for human and nonhuman primates (Flicker et al., 1987; Mishkin, 1978).

### Results

#### Traditional Memory Tests

We initially compared performances in the AD and AAMI groups on the standard memory tests noted above and recognized that by including variables that had been used in the original group definition we might obtain spuriously high rates of group discrimination. These variables did, however, allow for some intergroup overlap (i.e., lower levels of performance were not differentially restricted in the two groups). The only differential restrictions between groups on these measures were placed on the upper limits of performance.

A discriminant analysis with the traditional test variables block entered was significant, yielding an eigenvalue of 1.422 and a canonical correlation of .766 (Wilks's $\lambda = .413$, $\chi^2(4, N = 112) = 95.54$, $p < .0001$). It correctly classified 87.50% of all subjects. The final column of Table 1 presents the standardized canonical discriminant function coefficients for the individual test variables.

Inspection of Table 1 reveals that the greatest discriminant weight was associated with the BVRT and Logical Memory. The hard pairs from Paired Associate Learning contributed less to the discriminant function, whereas the easy paired associates contributed the least discriminant weight.

Table 2 presents the sensitivity and specificity (i.e., true−false positive and true−false negative rates) for this discriminant function. As can be seen, the more likely classification error was the misidentification of AAMI subjects as having AD.

#### Computer-Simulated Everyday-Memory Tests

The discriminant analysis with the computerized everyday-memory test variables block entered was also significant, yielding an eigenvalue of 1.584 and a canonical correlation of .783 (Wilks's $\lambda = .387$, $\chi^2(4, N = 112) = 102.53$, $p < .0001$). It correctly classified 88.39% of all subjects. The last column of Table 3 presents the standardized canonical discriminant function coefficients for the individual test variables. The greatest discriminant weight was given to recall of four Name−Face pairs. Misplaced Objects (total found) and Delayed Non-Matching-to-Sample (total correct) contributed less, with six Name−Face pairs contributing the least discriminant weight.

Table 4 presents the sensitivity and specificity for this discriminant function. A comparison of the hit rates of the traditional tests and the computer-simulated everyday-memory measures reveals that both batteries yielded generally equal rates of correct classification. The MAC battery committed one less false-positive error than did the traditional battery (i.e., the everyday-memory tests misidentified one fewer actual AAMI as AD than did the traditional tests). As with the traditional tests, the more frequent classification error was the misidentification of AAMI subjects as falling within the AD group.

### Table 4

<table>
<thead>
<tr>
<th>Classification Rates of Computer-Simulated Everyday Memory Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted group membership</td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Actual group membership</td>
</tr>
<tr>
<td>AD</td>
</tr>
<tr>
<td>AAMI</td>
</tr>
<tr>
<td>Total predicted cases</td>
</tr>
</tbody>
</table>

*Note. AD = Alzheimer's disease; AAMI = age-associated memory impairment.*
Discussion

This study demonstrated generally equal and reasonably high rates of accurate discrimination between mild AD and AAMI using batteries of either traditional psychometric instruments or computer-simulated everyday-memory tasks. This equivalence of accuracy occurred despite the use of the traditional measures in the original group definition, suggesting that the everyday-memory tests may be at least as accurate as the traditional measures in group discrimination. This is notable in that the everyday-memory tests were initially developed to provide measures of treatment outcome that had direct relevance to daily life and without discriminative validity in mind. The traditional tests, on the other hand, were first used in the differential diagnosis of organically caused cognitive dysfunction (i.e., discriminative validity), with little concern for ecological validity.

The importance of ecological validity and concepts of everyday memory has recently received strong support in the cognitive psychology literature (Ceci & Bronfenbrenner, 1991; Conway, 1991; Loftus, 1991; Morton, 1991; Neisser, 1991) after some earlier criticism (Banaji & Crowder, 1989). Given this emphasis, it appears to be a movement toward directly relevant, generalizable dependent variables in experimental neuropsychological research. This shift is becoming apparent in clinical neuropsychological assessment (e.g., California Verbal Learning Test, Delis, Kramer, Kaplan, & Ober, 1987; Rivermead Behavioural Memory Test, Wilson, Cockburn, Baddeley, & Hoirns, 1989). Our results suggest that it is possible to use dependent variables with more ecological and face validity in the study of mild-AD without sacrificing any discriminative validity.

The ability to accurately discriminate between AAMI and mild AD with either or both batteries has important clinical implications. Many otherwise healthy older individuals often correctly recognize that they have experienced significant memory decline relative to their younger years (i.e., AAMI). A number of these individuals suffer from fears that they are in the early stages of AD and eventually present in the clinic. Our results demonstrate that their normal cognitive decline can be differentiated from the pathological decline of mild AD with a reasonable amount of accuracy, allowing them to be reassured and their need for anxiety eliminated.

On the other hand, it is noteworthy that the greatest error of misclassification for both the traditional tests and the computer-simulated everyday memory measures was the false-positive error of misidentifying AAMI as AD. The original AAMI criteria (Crook et al., 1986) and subsequent revisions of these criteria (Blackford & La Rue, 1989) allow for the possibility that some subjects in the AAMI category may actually be in the early stages of AD. This factor may have contributed to our false-positive error rate, highlighting the need for longitudinal investigation of age-related cognitive decline.

References


Received April 4, 1991
Revision received June 21, 1991
Accepted June 21, 1991