Neuropsychology
2005, Vol. 19, No. 5, 629–640

Copyright 2005 by the American Psychological Association
0894-4105/05/$12.00 DOI: 10.1037/0894-4105.19.5.629

Early Assessment of Dementia: The Contribution of Different Memory Components

Pauline E. J. Spaan and Jeroen G. W. Raaijmakers
University of Amsterdam

Cees Jonker
Vrije Universiteit

A broad memory test battery (reflecting explicit and implicit memory functioning) was administered to a heterogeneous sample of initially nondemented, community-dwelling elderly subjects. To examine the profile of preclinical dementia, subjects were tested twice: At baseline, all subjects were nondemented according to *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) criteria; 2 years later, a subgroup had developed dementia. Performance of the preclinically demented subjects was best characterized, relative to that of cognitively impaired subjects who did not develop dementia 2 years later, by an inability to benefit at recall from semantic relations and by absent repetition priming effects. The authors conclude that in addition to testing episodic memory functioning, it is important to be aware of semantic and implicit memory deficits in the early assessment of dementia.

Keywords: preclinical dementia, mild cognitive impairment, episodic memory, semantic memory, priming

The difference between memory problems in dementia and purely age-associated memory impairments is subtle. Particularly at an early stage of dementia, the pathological effects of the disease are still rather modest, whereas in very old persons (i.e., 75 years and above), normal aging effects on cognitive performance may be more pronounced. In addition, the prevalence of dementia increases exponentially with increasing age (e.g., Ott et al., 1995). Research into the early predictors of dementia is, therefore, complicated, but also very important. The current medications aimed at altering the natural history of dementia are only useful when administered in the earliest possible stage. Thus, early assessment is important to the ability to offer appropriate (psycho)social care and support in time.

Therefore, the differences in memory performance between dementia patients and nondemented elderly controls are frequently investigated, both in experimental psychological studies and in clinical population-based studies. However, these two types of studies differ in several respects, which may complicate the search for the best differentiating memory variables. First of all, *experimental* studies usually investigate relatively young elderly subjects and patients (~68 years) with a high educational level (e.g., college education; ~12 years of attained education), comparing them in a cross-sectional research design (e.g., Spaan, Raaijmakers, & Jonker, 2003). *Clinical* studies, however, use (much larger) samples of subjects that are far more representative of the average (i.e., older and with a lower educational level) elderly population. These subjects are usually investigated longitudinally: At baseline, all subjects are nondemented, whereas at follow-up, a subgroup has developed dementia according to *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [*DSM–IV*]; American Psychiatric Association, 1994) clinical criteria. Consequently, these studies are able to investigate the preclinical stage of dementia, whereas experimental studies examine a more advanced stage. Recently, many studies have used a *mild cognitive impairment* (MCI) group of patients, who have subjective and objective memory impairments but intact functioning of activities of daily living (ADL; e.g., Petersen et al., 1999). It is assumed that relatively many MCI patients will convert to dementia, particularly Alzheimer's disease (AD), in the near future (e.g., Petersen et al., 2001). Therefore, MCI studies are also able to investigate the preclinical stage of dementia, at least an earlier stage than that examined in most experimental studies. It must be noted that the majority of MCI patients do not convert to dementia within 2 or 3 years, though the ratio is clearly higher than that based on age alone (e.g., Petersen et al., 2001).

Second, the types of memory tasks and/or the terminology used differ. Current experimental studies usually adopt a multicomponent approach in the conceptualization of memory. Generally, a division is made into explicit memory (discerning episodic and semantic memory) and implicit memory (e.g., reflecting priming and procedural memory; Schacter, 1992; Squire, 1992; Tulving, 1972). Furthermore, short-term (or working) memory is discerned, subdividing visuospatial and phonological–verbal information (Baddeley & Hitch, 1974). This multicomponent approach is not

Pauline E. J. Spaan and Jeroen G. W. Raaijmakers, Department of Psychonomics, University of Amsterdam, Amsterdam, the Netherlands; Cees Jonker, Institute for Research in Extramural Medicine and Department of Psychiatry, Vrije Universiteit, Amsterdam, the Netherlands.

Additional materials are on the Web at http://dx.doi.org/10.1037/0894-4105.19.5.629.supp

Part of this research was presented at the Third Joint Meeting of the Australian Society for the Study of Brain Impairment and the International Neuropsychological Society, Brisbane, Australia, July, 2004. This work was supported by Grant 970-10-031 from the Netherlands Organization for Scientific Research (NWO).

Correspondence concerning this article should be addressed to Pauline E. J. Spaan, University of Amsterdam, Department of Psychonomics, Roetersstraat 15, 1018 WB Amsterdam, the Netherlands. E-mail: p.e.j.spaan@uva.nl

clearly present within clinical, population-based studies or MCI studies that mainly use memory tests that are frequently administered in clinical practice. These tests make the tacit assumption that memory is a single entity, though it may be subdivided according to the modality of the material to be memorized (verbal, visual), the different demands on the reproduction process (free recall, cued recall, recognition), or the length of the interval between learning and reproduction phase (immediate recall, delayed recall). Nonetheless, in experimental memory terms, clinical memory testing is primarily based on the measurement of episodic memory, though some frequently used tests measure semantic memory as well.[1]

The multicomponent approach to memory is also absent in the MCI criteria (Petersen et al., 1999, 2001), which do not describe the types of memory impairments or tests to measure them. It is assumed that episodic memory deficits are required in MCI, but implementation is complicated by the numerous tests that measure the many forms of episodic memory (mentioned above).

Thus, it may be argued that clinical studies adopt a more global approach to memory testing in dementia, whereas experimental studies search for more specific causes and explanations of the presented deficits. Nonetheless, most experimental studies predominantly focus on one aspect of memory functioning instead of a broad range of aspects (which would enable observation of how these different memory aspects interact with one another).

A review of experimental studies on the specific memory disorders in dementia (AD in particular; e.g., Spaan et al., 2003) shows that in addition to episodic memory problems, there are also major semantic memory dysfunctions (e.g., Chertkow & Bub, 1990; Hodges & Patterson, 1995). AD patients exhibit, relative to unimpaired elderly controls, poor semantic encoding of to-be-learned information (e.g., Monsch et al., 1994; Randolph, Braun, Goldberg, & Chase, 1993). This will also affect episodic memory performance, especially with material that has an inherent semantic structure or in semantic cuing tasks: AD patients do not benefit from such cues, unlike unimpaired elderly controls (e.g., Bird & Luszcz, 1991; Monti et al., 1996). AD patients might not be able to discriminate between two related concepts, because the attribute knowledge that distinguishes the two concepts has been lost (e.g., Martin & Fedio, 1983; Sailor, Bramwell, & Griesing, 1998). In addition, AD patients' deficits are evident in recognition tasks, particularly when semantically related distractors are used—their responses contain many false-positive errors (e.g., Brandt, Corwin, & Krafft, 1992; Deweer et al., 1994). AD patients seem unable to inhibit irrelevant associations (e.g., Helkala, Laulumaa, Soininen, & Riekkinen, 1989). In addition, results of priming experiments based on more conceptually (i.e., semantically) based encoding tasks reflect deficits in AD patients' performance as well, once again because of their impaired semantic capacities (e.g., Gabrieli et al., 1994; Keane, Gabrieli, Fennema, Growdon, & Corkin, 1991; Meiran & Jelicic, 1995). Also, AD patients' poor visuospatial span, relative to their auditory–verbal span, has been reported frequently (e.g., Carlesimo, Fadda, Lorusso, & Caltagirone, 1994; Trojano, Chiacchio, DeLuca, & Grossi, 1994).

A review of available clinical, longitudinally based studies on predictors of dementia (e.g., Collie & Maruff, 2000; Spaan et al., 2003) shows that deficits on the following tests may be indicative of developing AD, several years before the diagnosis is made: (delayed) story recall, similarities (WAIS–R), (verbal) paired-associate learning, (delayed) free recall (and recognition) of words, immediate visual memory, digit symbol (WAIS–R), and verbal fluency. From these data and from a review of experimental studies that examined patients in their earliest stages ("minimal" AD) or diagnosed AD patients with high scores on cognitive screening tests (i.e., Mini-Mental State Examination [MMSE] scores greater than 23; Folstein, Folstein, & McHugh, 1975), we conclude that tests sensitive to semantic knowledge are crucial for detecting AD at the earliest possible stage. These tests may include memorizing material with an inherent semantic structure (e.g., story recall), semantic cuing (as in verbal paired-associate learning), mental lexicon, verbal abstract reasoning (e.g., similarities), or category fluency. It is possible that reliable priming tasks that call on semantic processing may also be useful (Spaan et al., 2003).

In the present study, we investigated whether the differentiation between various memory components (i.e., episodic, semantic, implicit, working memory) contributes to the early assessment of dementia. Memory performance of a heterogeneous sample (in terms of age, education, and various cognitive functions) of initially nondemented, community-dwelling elderly subjects was investigated. At a first baseline measurement (T_1), none of the subjects was clinically demented; at a second measurement 2 years later (T_2), a subgroup was officially diagnosed as demented (according to *DSM–IV* criteria).

In the first section below, we determine the combination of measures (i.e., memory components) that is most accurate in predicting dementia before the diagnosis can officially be made (i.e., at T_1). It is expected that tests sensitive to semantic processing capacities should be most sensitive to preclinical dementia (PCD). Additional analyses were performed over a subgroup of elderly subjects who were defined as *cognitively impaired* (CI) on the basis of MMSE scores and tests of delayed verbal recall, processing speed, and nonverbal abstract reasoning ability. These CI subjects were elderly subjects with cognitive deficits, comparable to the MCI concept (more specifically, the *multiple-cognitive-deficits* type), and they represented a sample that was obviously more relevant for clinical practice. In performing these analyses, we determined which variables are sufficiently sensitive to differentiate PCD cases from CI subjects who did not convert to dementia at T_2. In the second section, we present cross-sectional analyses of the measures that best discriminate between demented and nondemented subjects at the time of diagnosis (T_2), and we investigate if and how memory performance characteristics of the PCD and the clinically demented subjects differed.

[1] Note that there are several clinical tests that do measure aspects of semantic memory (e.g., several verbal subtests of the Wechsler Adult Intelligence Scale—Revised [WAIS–R; Wechsler, 1981], the Boston Naming Test [Kaplan, Goodglass, & Weintraub, 1983], tests of category fluency). However, these tests are not always interpreted as measures of semantic memory but, instead, are interpreted as measures of verbal intelligence or language functioning. Fortunately, nowadays there is increasing awareness of the semantic nature of these tests in neuropsychological assessment. However, tests of implicit memory functioning are (still) absent in clinical assessment.

Method[2]

Participants

Subjects were selected from the Longitudinal Aging Study Amsterdam (LASA), a large-scale, population-based study of elderly individuals (Deeg, Beekman, Kriegsman & Westendorp-de Serière, 1998). At T_1, recent LASA data were used to select subjects for the current research (see Appendix B, which is available on the Web at http://dx.doi.org/10.1037/0894-4105.19.5.629.supp). All subjects were community dwelling and had been screened for depression (Center for Epidemiological Studies Depression scale [CES–D]; Beekman et al., 1997; Radloff, 1977), a history of cerebrovascular accidents (CVAs), prevalent dementia, and other neurologic or psychiatric causes of cognitive dysfunctioning (Cambridge Examination for Mental Disorders of the Elderly [CAMDEX], meeting *DSM–IV* criteria; Neri, Rubichi, DeVreese, Roth, & Cipolli, 1998).

In addition, to create a cognitively heterogeneous sample at T_1 (which should result, at T_2, in as many demented subjects as possible), we created two subgroups according to their global level of cognitive functioning, measured by the MMSE (Folstein et al., 1975): the *CI* group and the *normal control* (NC) group. The CI group scored in the 21–25 range of the MMSE. The NC group scored in the 27–30 range of the MMSE (see Tombaugh and McIntyre, 1992 and Footnote 1 of Appendix A, which is available on the Web at http://dx.doi.org/10.1037/0894-4105.19.5.629.supp, for justification). CI subjects were matched to NC subjects for age, years of education, and sex.

Additional data derived from recent LASA data showed that the subjects screened for the CI group performed significantly worse than the NC subjects on three available cognitive measures: the 15 Words Test (after Rey, 1964), the Coding Task (adapted version of a letter substitution task, the Alphabet Coding Task–15; Savage, 1984), and the Raven Colored Progressive Matrices (after Raven, 1984). Details on the versions of the tasks that were used are described in Deeg and Westendorp-de Serière (1994). The CI group was supposed to be at risk for developing dementia (within a few years). The NC group was supposed to be cognitively intact. The CI subjects might be interpreted as suffering from MCI (Petersen et al., 1999). The one criterion that we could not meet in this study is the presence of subjective memory impairments, because we did not have systematic information on this criterion (although subjective memory complaints were frequently voiced when subjects were initially tested). Thus, the CI group did not officially meet the MCI criteria, though there was a high resemblance, most likely to the multiple-cognitive-deficits type (e.g., Busse, Bischkopf, Riedel-Heller, & Angermeyer, 2003; Lopez, Jagust, DeKosky, et al., 2003; Lopez, Jagust, Dulberg, et al., 2003). Therefore, we apply the descriptive term *CI* rather than MCI, which is associated with Petersen's criteria. Ultimately, at T_1, valid and complete data sets were available for 119 clinically nondemented subjects: 51 CI subjects and 68 NC subjects.

These subjects were approached for a second test 2 years later (T_2). Valid and complete data sets were available for 96 subjects at T_2 (43 CI subjects and 53 NC subjects). In addition, we readministered the CAMDEX to determine which subjects had developed dementia in the intervening period (i.e., CAMDEX diagnosis of *mild/moderate dementia*; $n = 6$). Furthermore, the general practitioners of the subjects who dropped out at T_2 were asked whether dementia was assessed according to the official medical records (resulting in 3 additional demented cases). All of these 9 subjects who were demented at T_2 originated from the CI group created at T_1. In this way, three clinical subgroups were obtained, based on the dementia assessment at T_2 and screening of cognitive performance at T_1:

1. *NC* ($n = 68$): no cognitive dysfunctions; intact ADL functioning; not demented; not depressed; no CVA history.
2. *CI* ($n = 42$): various cognitive dysfunctions (i.e., MMSE, delayed recall, processing speed, nonverbal abstract reasoning ability), but intact ADL functioning and not demented; not depressed; no CVA history. These subjects were considered to be at risk for developing dementia in the (near) future.
3. *PCD* ($n = 9$): various cognitive dysfunctions; intact ADL functioning and not demented at T_1, but diagnosed as demented according to *DSM–IV* criteria at T_2—preclinically demented at T_1; clinically demented at T_2 (most likely cause: AD, considering the absence of CVA history or depression).

Table 1 presents further characteristics (demographic and screening variables) of these three clinical subgroups. The subjects that were retested at T_2 ($n = 96$), and the subjects who dropped out at T_2 ($n = 23$) differed significantly only in age (mean age at T_1 of dropped-out subjects: 82.00 years [$SD = 6.58$]; mean age at T_2 of retested subjects: 78.10 years [$SD = 8.45$]), $p = .008$.

General Procedure

The administration of the memory test battery took place in the home environment of the subject (after informed consent) by means of a laptop computer, which was operated by a neuropsychologist (Pauline E. J. Spaan); the subject only had to look at the screen. Within 6 weeks of the test administration, the CI subjects were administered the short version of the CAMDEX (Neri et al., 1998) by a physician. The procedure of the repeated memory test administration at T_2, 2 years later, was identical to the procedure at T_1, except for the additional administration of the CES–D and questions regarding CVA to control for those variables (as was done at T_1). In addition, subjects with MMSE scores less than 27 at T_2 were administered the CAMDEX by the physician to identify the incident demented cases. All subjects were discussed with an experienced neurologist (Cees Jonker), who was responsible for the final assessment. Because it was not possible to conduct a physical examination, the diagnosis simply said *demented* or *not demented* rather than specifying a type of dementia.[3]

Materials

Memory Test Battery

The memory tests were administered by means of an Apple PowerBook laptop computer (Cupertino, CA). All stimuli were presented onscreen.

[2] More detailed methodological information (regarding subjects, general procedures, and the subtests of the memory test battery) is provided in Appendix A, which is available on the Web at http://dx.doi.org/10.1037/0894-4105.19.5.629.supp.

[3] The CAMDEX does specify the severity of the disease: *minimal*, *mild*, *moderate*, or *severe* dementia. In mild, moderate, and severe dementia, all of the *DSM–IV* criteria of dementia are satisfied (i.e., the PCD group in the current research). In minimal dementia, the *DSM–IV* criteria are not satisfied, because (a) no additional cognitive deficits are observed (only ["episodic"] memory impairments are found), and (b) cognitive deficits do not cause evident impairment in occupational or social functioning. Therefore, subjects with a minimal dementia diagnosis were classified as *nondemented* (all originated from the CI group; see Roth et al., 1986, for more detailed information). It may be argued that the subjects with a minimal dementia classification meet criteria for the MCI *amnestic* type (e.g., Busse et al., 2003; Lopez, Jagust, DeKosky, et al., 2003; Lopez, Jagust, Dulberg, et al., 2003; Luis et al., 2003). These subjects are, however, not presented as a separate group, because the category seems rather unstable regarding conversion to dementia (for details, see Appendix B, which is available on the Web at http://dx.doi.org/10.1037/0894-4105.19.5.629.supp).

Table 1
Average Demographic Characteristic and Screening Variable Data (Means, With Standard Deviations in Parentheses) for Each Clinically Relevant Subgroup of Subjects

Variable	NC (n = 68): T_1 MMSE > 26	CI (n = 42): T_1 MMSE < 26	PCD (n = 9): T_2-demented; T_1 MMSE < 26	NC + CI (n = 110): All nondemented	CI + PCD (n = 51): T_1 MMSE < 26
Age[a,b]	78.04 (7.42)	80.14 (8.53)	87.00 (3.57)	78.85 (7.89)**	81.35 (8.29)*
Education (years)	7.57 (2.65)	7.48 (1.99)	8.44 (3.13)	7.54 (2.41)	7.65 (2.23)
Sex (no. male/no. female)	34/34	23/19	4/5	57/53	27/24
MMSE[a,c]	28.07 (1.04)	23.64 (1.91)	22.89 (1.45)	26.38 (2.59)**	23.51 (1.85)**
CES–D	6.29 (3.66)	6.05 (4.77)	9.11 (5.35)	6.20 (4.10)*	6.59 (4.96)
15 Words Test: Delayed recall[a,b,c]	5.63 (2.57)	3.96 (2.22)	0.57 (0.79)	5.06 (2.56)**	3.29 (2.43)**
Coding Task[a,c]	67.90 (12.15)	51.94 (14.01)	49.80 (13.85)	62.17 (14.91)	51.66 (13.82)**
Raven CPM[a,c]	17.66 (3.40)	14.44 (3.74)	13.33 (2.69)	16.45 (3.85)*	14.24 (3.57)**

Note. Data represent analysis of variance results with Bonferroni post hoc analyses over normal control (NC), cognitively impaired (CI), and preclinically demented (PCD) subgroups. Independent-samples t tests indicated significant differences between the NC + CI and PCD groups and significant differences between the NC and CI + PCD groups. T_1 = time of first baseline measurement; MMSE = Mini-Mental State Examination (Folstein et al., 1975); T_2 = time of second measurement (2 years after T_1); no. = number; CES–D = Center for Epidemiological Studies Depression scale (Beekman et al., 1997; Radloff, 1977); CPM = Colored Progressive Matrices.
[a] Significant difference ($p < .05$) between the NC and PCD groups. [b] Significant difference ($p < .05$) between the CI and PCD groups. [c] Significant difference ($p < .05$) between the NC and CI groups.
* $p < .05$. ** $p < .01$.

Below, the subtests are briefly described in the same order as they were presented in the battery.

10-Word List-Learning Test (episodic memory). This consisted of free recall of 10 semantically unrelated words in three trials. Between presentation and recall phase, the subject performed a distraction task for 20 s (to prevent recency effects). Main score in data analyses: total number of words reproduced over three trials (range: 0–30).

Digit Span Task (working memory). This consisted of 10 trials of oral reproduction of a sequence of digits in the same order as they were presented. After a correct response, the next sequence was one digit longer and vice versa for a wrong response. A span score was computed that was best representative of performance (i.e., was least susceptible to chance).

Word-Recognition Test (yes–no; episodic memory). This involved explicit recognition ("yes" or "no") of the words of the 10-Word List-Learning Test from a list also including 10 semantically related distractors. Main score: sum of true-positive and true-negative answers (range: 0–20).

Paired-Associate Learning Test (episodic–semantic memory). This involved cued recall of five semantically related and five semantically unrelated word pairs (in the same format as the 10-Word List-Learning Test). Main score: sum of pairs reproduced over three trials (range: 0–30).

Block Span Task (working memory). This was a visuospatial variant of the Digit Span Task. Ten square fields ("blocks") were presented onscreen, like a chessboard. A sequence of random blocks flashed consecutively; the subject had to indicate the correct blocks in the correct sequence.

Word-Stem Completion Task (implicit memory). Two- or three-letter word stems had to be completed with the first (Dutch) word that came to mind. The test consisted of 10 *experimental* stems—which could be completed with the words from the 10-Word List-Learning Test—and 10 *control* stems. The subject was not alerted to the connection with the 10-Word List-Learning Test. Main score: number of experimental stems completed with a target word minus the number of control stems completed with a target word (range: −10–10).

Category Fluency Test (semantic memory). Subjects had to name as many exemplars that belonged to the categories of "animals" and "occupations" as they could think of within 60 s per category. Main score: sum of correct answers over both categories.

Mirror-Reading Task (implicit memory). Sixty mirror words (rotated on the vertical axis) had to be read as quickly as possible. Main score: mean reading time (in seconds) over all correctly read mirror words.

Perceptual Identification Task (semantic–implicit memory). The subject had to read words that were briefly presented onscreen as quickly as possible. Each word was repeatedly presented in *tics* of 16 ms: Each consecutive presentation lasted one tic longer, alternated by a mask. The test consisted of 48 words: 12 *low-frequency* (LF), 12 *middle-frequency* (MF), 12 *high-frequency* (HF), and 12 *repeated* presentations of the MF words (rep-MF; to measure repetition priming effects). The subject was not alerted to the repeated presentation. Main *semantic memory* score: mean reaction time (RT) over all HF, MF, and LF words. Main *priming* score: mean RT for the MF words minus the mean RT for the rep-MF words.

Two-Alternative Word-Recognition Test (episodic memory). This involved explicit recognition of the words from the 10-Word List-Learning Test while each word's semantically related distractor was simultaneously presented. Main score: sum of correctly recognized words (range: 0–10).

Visual Association Test (episodic memory). This involved cued recall of six line drawings of common objects that had been previously presented in interaction with another object or cue, representing an illogical combination (Lindeboom, Schmand, Tulner, Walstra, & Jonker, 2002). Main score range: 0–6.

The Profile of Memory Measures Best Predicting Dementia Before the Diagnosis Can Be Made (T_1)

In this section, we investigate how the subjects who were diagnosed as clinically nondemented at T_1 but turned out to be demented at T_2 (according to CAMDEX and *DSM–IV* criteria; the PCD subjects) can be differentiated from the initially and presently nondemented subjects. In other words, what are the specific performance characteristics of PCD subjects, or which profile (combination) of memory measures is best predictive of dementia 2 years in advance? And how accurate is this prediction? Does this profile lead to an improvement in the prediction of dementia above what is possible using the current clinical assessment methods (e.g., the MMSE, purely episodic memory tests)?

Table 2
Average Main Scores (Means, With Standard Deviations in Parentheses) per Memory Test Battery Subtest, Administered at T_1, for Each Clinically Relevant Subgroup of Subjects

Subtest (Cronbach's alpha)	NC (n = 68): T_1 MMSE > 26	CI (n = 42): T_1 MMSE < 26	PCD (n = 9): T_2 demented; T_1 MMSE < 26	NC + CI (n = 110): All nondemented
10-Word List-Learning Test (α = .72)[a,b,c]	13.56 (4.46)	10.48 (4.19)	6.22 (2.77)	12.38 (4.59)**
Digit Span Task[a,c]	4.60 (0.98)	4.07 (0.81)	4.89 (0.60)	4.40 (0.95)*
Word-Recognition Test (α = .66)[a,b,c]	18.41 (1.73)	17.36 (2.35)	14.89 (2.67)	18.01 (2.04)**
Paired-Associate Learning Test (α = .87)[a,b,c]	18.26 (5.50)	13.40 (4.84)	4.22 (5.04)	16.41 (5.75)**
Block Span Task[a]	3.99 (0.99)	3.43 (0.74)	3.33 (0.71)	3.77 (0.94)
Word-Stem Completion Task (α = .26)	0.38 (0.57)	0.31 (0.52)	0.33 (0.50)	0.35 (0.55)
Category Fluency Test (α = .73)[a,b]	26.46 (7.02)	22.10 (5.63)	19.22 (7.17)	24.79 (6.84)*
Mirror-Reading Task (α = .84)[a]	4.76 (2.03)	7.84 (3.73)	6.86 (5.42)	5.94 (3.16)
Perceptual Identification Task				
Semantic memory (α = .97)[a]	6.08 (1.48)	7.84 (2.84)	7.55 (1.47)	6.75 (2.26)
Priming (α = .31)[a,c]	0.51 (0.62)	1.00 (0.86)	0.12 (0.86)	0.70 (0.76)*
Two-Alternative Word-Recognition Test (α = .67)[a,b,c]	9.46 (0.97)	8.71 (1.69)	6.78 (1.99)	9.17 (1.33)**
Visual Association Test (α = .75)[a,b,c]	5.49 (0.70)	4.95 (1.23)	3.11 (1.83)	5.28 (0.97)**

Note. Data represent analysis of variance results with Bonferroni post hoc analyses over normal control (NC), cognitively impaired (CI), and preclinically demented (PCD) subgroups. Independent-samples t tests indicated significant differences between the NC + CI and PCD groups. T_1 = time of first baseline measurement; MMSE = Mini-Mental State Examination (Folstein et al., 1975); T_2 = time of second measurement (2 years after T_1).
[a] Significant difference ($p < .05$) between the NC and CI groups. [b] Significant difference ($p < .05$) between the NC and PCD groups. [c] Significant difference ($p < .05$) between the CI and PCD groups.
* $p < .05$. ** $p < .01$.

First, we investigated how and how well the memory test battery and other variables predicted dementia within the entire group of subjects who were initially (at T_1) nondemented (n = 119; i.e., including subjects with high MMSE scores, the NC group). Additional analyses were focused on the subgroup of subjects who were originally classified to the CI group (i.e., the subjects assumed to be at risk for developing dementia because of their cognitive dysfunctions; n = 51 [9 of whom converted to dementia at T_2: the PCD subjects]). Table 2 presents an overview of mean scores per subtest for the nondemented and the PCD subjects, both within the entire group of subjects and within the CI group.

The Prediction of Dementia in the Entire Group of Elderly Subjects

The upper half of Table 3 shows the results of several discriminant analyses performed over the T_1 scores per subtest and several subject-related variables (i.e., age, education, sex, MMSE, and CES–D). First, all of these variables or meaningful clusters of variables (e.g., clinically used variables such as the MMSE or typical episodic memory tests[4]) were entered together. In this way, the accuracy of these variables was examined to classify the subjects to the nondemented or to the PCD group. In addition, the characteristics of this classification process were described by means of the percentages of *true-positive* classifications (i.e., subjects diagnosed as PCD being classified as PCD; *sensitivity*) and *true-negative* classifications (i.e., nondemented subjects being classified as nondemented; *specificity*). Furthermore, a discrimination measure (d') was calculated as a function of the true-positive rate and the false-positive rate.[5] Second, we investigated by means of a stepwise analysis, how these measures may be reduced to create a selection of variables that were best able to discriminate between the PCD subjects and the nondemented subjects. Classification was performed with an equal probability (50%) for subjects being classified to the PCD or to the nondemented group (Method A), as well as with the prior probability being computed from the respective group size (Method B; see Table 3).

First, it may be noted that the stepwise discriminant analysis did not select the T_1 MMSE score as one of the best discriminating variables. Only 78.2% of the originally grouped subjects were correctly classified using the MMSE (according to Method A), relative to 90.8% using the best discriminating combination of the Paired-Associate Learning Test, the Visual Association Test, and the priming measure of the Perceptual Identification Task. Also, note the large difference in d' between the two analyses (1.54 vs. 2.58; see Table 3). The T_1 MMSE score falsely classified many subjects to the PCD group (many false positives, resulting in a worse specificity: 78%). In addition, the sensitivity of the MMSE (78%) was worse than the sensitivity of the three best discriminating variables (89.0%). These two characteristics of the MMSE explain the low d' value.

The purely episodic memory measures evidently differentiated better between the PCD subjects and the nondemented subjects. Indeed, one of the episodic memory measures (i.e., the Visual Association Test) was selected by the stepwise analysis as best discriminating variable. However, the Paired-Associate Learning

[4] These include the 10-Word List-Learning Test, the Word-Recognition Test, the Two-Alternative Word-Recognition Test, and the Visual Association Test. The Paired-Associate Learning Test is not included in the cluster of episodic memory measures because semantic memory functioning is considered to be important in this subtest as well (i.e., an impaired performance in the demented subjects is mainly characterized by not being able to benefit from semantic relations between words, as is discussed and illustrated below; see Figure 1 [presented later]).

[5] The true-positive rate (or sensitivity) measure was calculated as follows: true-positive predictions/number of subjects diagnosed as demented. The false-positive rate (or "inverted specificity") measure was calculated as follows: false-positive predictions/number of subjects diagnosed as nondemented.

Table 3
Predictions of Dementia (T_2 CAMDEX: Mild or Moderate Dementia: PCD Cases) According to Various Variables Measured at T_1

Variable	Classification Method[a]	Classification Accuracy	Percentage True positives	Percentage True negatives	d'
NC + CI (n = 110) versus PCD (n = 9)					
None (baseline): "Blind" classification on the basis of chance	A	50.0%	4.5/9 (50%)	55/110 (50%)	0.00
	B_1	86.6%	1/9 (11%)	102/110 (93%)	0.24
All variables together: All memory measures and age, education, sex, T_1 MMSE score, and T_1 CES–D score	A	97.5%	8/9 (89%)	108/110 (98%)	3.28
	B_1	97.5%	7/9 (78%)	109/110 (99%)	3.09
All variables, stepwise: Paired-Associate Learning Test (residual variance = .467), Visual Association Test (residual variance = .360), and Perceptual Identification Task (priming; residual variance = .314)	A	90.8%	8/9 (89%)	100/110 (91%)	2.58
	B_1	95.0%	6/9 (67%)	107/110 (97%)	2.32
T_1 MMSE score	A	78.2%	7/9 (78%)	86/110 (78%)	1.54
T_1 Visual Association Test	A	81.5%	6/9 (67%)	91/110 (83%)	1.39
T_1 Paired-Associate Learning Test	A	81.5%	8/9 (89%)	89/110 (81%)	2.11
Purely episodic memory measures together: 10-Word List-Learning Test, Word-Recognition Test, Two-Alternative Word-Recognition Test, and Visual Association Test	A	89.9%	7/9 (78%)	100/110 (91%)	2.11
	B_1	94.1%	4/9 (44%)	108/110 (98%)	1.90
CI (n = 42) versus PCD (n = 9)					
None (baseline): "Blind" classification on the basis of chance	A	50.0%	4.5/9 (50%)	21/42 (50%)	0.00
	B_2	72.5%	2/9 (22%)	35/42 (83%)	0.18
All variables together: All memory measures and age, education, sex, T_1 MMSE score, and T_1 CES–D score	A	94.1%	9/9 (100%)	39/42 (93%)	3.79
	B_2	94.1%	8/9 (89%)	40/42 (95%)	2.87
All variables, stepwise: Paired-Associate Learning Test (residual variance = .530), Perceptual Identification Task (priming; residual variance = .464), and Visual Association Test (residual variance = .414)	A	86.3%	7/9 (78%)	37/42 (88%)	1.94
	B_2	94.1%	7/9 (78%)	41/42 (98%)	2.82
T_1 MMSE score	A	60.8%	6/9 (67%)	25/42 (60%)	0.70
T_1 Visual Association Test	A	68.6%	6/9 (67%)	29/42 (69%)	0.94
T_1 Paired-Associate Learning Test	A	86.3%	7/9 (78%)	37/42 (88%)	1.94
Purely episodic memory measures together: 10-Word List-Learning Test, Word-Recognition Test, Two-Alternative Word-Recognition Test, and Visual Association Test	A	78.4%	7/9 (78%)	33/42 (79%)	1.58
	B_2	84.3%	3/9 (33%)	40/42 (95%)	1.20

Note. T_2 = time of second measurement (2 years after T_1); CAMDEX = Cambridge Examination for Mental Disorders of the Elderly (Neri et al., 1998); T_1 = time of first baseline measurement; NC = normal control; CI = cognitively impaired; PCD = preclinically demented; MMSE = Mini-Mental State Examination (Folstein et al., 1975); CES-D = Center for Epidemiological Studies Depression scale (Beekman et al., 1997; Radloff, 1977).
[a] A = all groups equal probability (50%); B_1 = prior probability computed from group size (nondemented group: 110/119 → 92.4%; PCD group: 9/119 → 7.6%); B_2 = prior probability computed from group size (nondemented group: 42/51 → 82.4%; PCD group: 9/51 → 17.6%).

Test and the priming measure of the Perceptual Identification Task apparently added more to the most efficient prediction of the PCD cases than did the other episodic memory measures (i.e., the 10-Word List-Learning Test, the Word-Recognition Test, and the Two-Alternative Word-Recognition Test).

Thus, poor performances on, in particular, the Paired-Associate Learning Test, the Visual Association Test, and the priming measure of the Perceptual Identification Task were key variables in predicting the PCD cases within a heterogeneous sample of elderly subjects (see Table 2 for a comparison of average values of these variables between the two groups). This selection of variables leads to a prediction accuracy that is only slightly lower than the accuracy of all variables together (see Table 3; sensitivity is equally high, but specificity is worse). According to Cohen's kappa analyses, the classification results of these three best discriminating measures are valid (despite the small number of PCD subjects [n = 9] relative to the number of nondemented subjects [n = 110]), and they are significantly better than classification on the basis of mere chance (see Table 4; see Lowry, 2005, for useful software).

The Prediction of Dementia in the CI Group

In this section, we describe additional analyses that were performed over the subgroup of subjects who were originally (at T_1) classified to the CI group (i.e., the *dementia-at-risk* group). It may be noted from the lower half of Table 3 that the stepwise analysis selected the same best discriminating measures within the CI group as within the entire group of subjects, though in a slightly different order of significance. Within the CI group, the priming measure of the Perceptual Identification Task made a more significant contribution than did the Visual Association Test, an episodic measure, despite the low level of reliability of the priming measure (α = .31; see Table 2). Thus, worse performance at T_1 on the Paired-Associate Learning Test, smaller repetition priming effects in the Perceptual Identification Task, and worse cued recall performance on the Visual Association Test predicted dementia best within a dementia-at-risk group 2 years before the diagnosis was made. The classification results of these three best discriminating measures were valid and significantly better than classification on

Table 4
Cohen's Kappa Analyses Performed Over Classification Results According to Best Discriminating Variables Between Preclinically Demented and Nondemented Subjects at T_1

Best discriminating variables[a]	Class. method[b]	Observed kappa (*SE*)	Proportion of agreement		
			Max. possible	Chance expected	Observed (.95 conf. interval)
Within entire group (n = 119)	A	.55 (.13)	.92	.80	.91 (.84–.95)
	B_1	.64 (.14)	1.00	.86	.95 (.89–.98)
Within CI group (n = 51)	A	.58 (.15)	.94	.67	.86 (.73–.94)
	B_2	.79 (.12)	.98	.72	.94 (.83–.98)

Note. T_1 = time of first baseline measurement; Class. = classification; Max. = maximum; conf. = confidence; CI = cognitively impaired.
[a] Paired-Associate Learning Test, Visual Association Test, and Perceptual Identification Task Priming. [b] A = all groups equal probability (50%); B_1 = prior probability computed from group size (nondemented group: 110/119 → 92.4%; preclinically demented [PCD] group: 9/119 → 7.6%); B_2 = prior probability computed from group size (nondemented group: 42/51 → 82.4%; PCD group: 9/51 → 17.6%).

the basis of mere chance (see Table 4). Table 2 presents a comparison of average values of these variables between the two groups.

Furthermore, the MMSE, again, had an evidently worse accuracy of classification, as well as a lower level of d', than the three best discriminating variables (as indicated by the stepwise analysis; see Table 3). The MMSE, using Method A, led to an accuracy only slightly above baseline. Once again, the purely episodic memory measures clearly discriminated better between the two groups than did the MMSE, though they were clearly worse than the Paired-Associate Learning Test, the priming measure of the Perceptual Identification Task, and the Visual Association Test together (i.e., they had worse specificity).[6]

Characteristics of Performance on Best Predicting Variables

The results described above show that the Paired-Associate Learning Test, the priming measure of the Perceptual Identification Task, and the Visual Association Test are the most sensitive predictors of subsequent dementia. These three measures differentiated well even within a group of elderly subjects with cognitive deficits (the CI group). To further examine the characteristics of performance of the PCD subjects, Figures 1 and 2 illustrate the patterns of performance across different trials and conditions in the first two tasks (the Visual Association Test does not consist of different trials or conditions). More specifically, we investigated whether PCD subjects benefited less (or not at all) from the conditions of semantic relatedness and implicit repetition of words (measuring semantic and implicit memory, respectively). Performance levels of the CI group and the PCD group are presented; within the entire group of subjects (also including the NC subjects), differences between the PCD subjects and nondemented subjects were even greater.

General linear model (GLM) repeated measures analyses of variance (ANOVAs) performed over the Paired-Associate Learning Test found significant Trial × Diagnostic Group interactions for the semantic pairs, $F(2, 48) = 5.081$, $p = .010$,[7] and the nonsemantic pairs, $F(2, 48) = 4.050$, $p = .024$. As illustrated by Figure 1, in both conditions, the CI subjects improved their recall performance over trials, whereas the PCD subjects showed an almost flat learning curve. A significant Trial × Diagnostic Group interaction was also found irrespective of condition, $F(2, 48) = 8.307$, $p = .001$. In addition, a significant Condition (semantic or nonsemantic pairs) × Diagnostic Group interaction, $F(1, 49) = 9.154$, $p = .004$, shows that the CI subjects benefited to a greater degree from the semantic relations within the material to be learned than did the PCD subjects.

GLM repeated measures ANOVAs performed over the Perceptual Identification Task found a significant effect of within-subject factor condition (MF vs. rep-MF words), $F(1, 49) = 12.701$, $p = .001$, but no significant effect of the between-subjects factor diagnostic group, $F(1, 49) = 0.069$, $p = .794$. In addition, a significant Condition × Diagnostic Group interaction was found, $F(1, 49) = 7.921$, $p = .007$. Thus, as Figure 2 illustrates, the CI subjects benefited from the repetition of words, whereas the PCD subjects did not. This difference in priming effect is not affected by the slower initial identification times for the nondemented subjects compared with the PCD subjects, because no significant overall group difference in identification times was found.[8]

The Profile of Memory Measures Best Discriminating Between Demented and Nondemented Elderly Subjects at the Time of Diagnosis (T_2)

The previous section focused on the prediction of dementia before the diagnosis could be made. For reasons of comparison, it

[6] An exploratory analysis showed that when the 5 subjects who were classified as minimally demented by CAMDEX at T_2 (but were nondemented at T_1—see Footnote 3 and Appendix B, which is available on the Web at http://dx.doi.org/10.1037/0894-4105.19.5.629.supp) were classified to the PCD group (instead of to the CI group) in the stepwise discriminant analysis, again, the Paired-Associate Learning Test and the priming measure of the Perceptual Identification Task were the best differentiating measures. In this analysis, however, the Visual Association Test was not selected as a significant predictor.

[7] In the semantic condition, the sphericity assumption was not met (Mauchly's test of sphericity: $p = .011$, ε [Greenhouse–Geisser] = .854). An adjusted test showed a significant Trial × Diagnostic Group interaction, $F(1.707, 83.662) = 7.297$, $p = .002$.

[8] In addition, analysis of this task within the entire group of subjects also showed a significant improvement in RTs in the nondemented group (whereas the PCD group did not improve): significant Condition × Diagnostic Group interaction, $F(1, 117) = 4.837$, $p = .030$. This effect occurred despite the shorter RTs of the nondemented subjects in general (though not significantly), $F(1, 117) = 2.682$, $p = .104$.

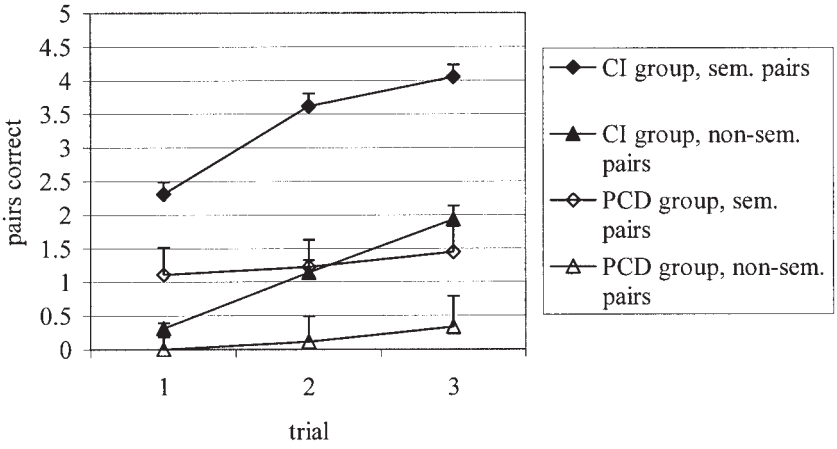


Figure 1. T_1 (i.e., baseline) performance on the Paired-Associate Learning Test, per trial and condition (semantic [sem.] vs. nonsemantic [non-sem.] pairs), by the cognitively impaired (CI) group versus the preclinically demented (PCD) group. Error bars represent standard error of the mean.

is also interesting to examine which measures of the memory test battery discriminated best between demented and nondemented subjects at the time of diagnosis (i.e., a cross-sectional analysis). Therefore, we performed additional analyses using the T_2 memory subtest data (instead of the T_1 data).

A stepwise discriminant analysis showed that within the entire group of subjects, the Visual Association Test (residual variance = .309) and the Word-Recognition Test (residual variance = .238) discriminated best (equal prior probabilities: accuracy of classification = 94.8%, d' = 2.70). However, within the CI group, the Word-Recognition Test (residual variance = .397) and the Category Fluency Test (residual variance = .291) were the best discriminating variables (equal prior probabilities: accuracy of classification = 90.7%, d' = 3.55). These latter two measures discriminated particularly well within the CI group ($d' > 3.00$).

It should be emphasized that these measures are frequently applied (in a slightly modified form) in clinical practice. Our data thus confirm the differentiating value of these measures once dementia can be clinically assessed. However, the predictive value of these measures in PCD and their sensitivity within a group of CI elderly persons is less impressive. Thus, for example, the Word-Recognition Test did not differentiate well between PCD subjects and nondemented subjects, as was shown by the T_1 data (CI mean score = 17.36 [*SD* = 2.35] vs. PCD mean score = 14.89 [*SD* = 2.67], on average 1 standard deviation below CI performance level). However, this subtest did discriminate well at T_2, when the diagnosis was really made (CI mean score = 17.59 [*SD* = 1.83] vs. PCD mean score = 13.00 [*SD* = 2.12], on average 2.5 standard deviations below CI performance level). The same was found for the Category Fluency Test: At T_1, the PCD group scored 0.5 standard deviations below the CI group's performance level, whereas the performance difference at T_2 had increased to 2.1 standard deviations (CI mean score = 22.70 [*SD* = 5.46] vs. PCD mean score = 11.00 [*SD* = 4.69]).

Conclusion

The most important finding presented in this article is the strong predictive value of, first, the Paired-Associate Learning Test; second, the priming measure of the Perceptual Identification Task;

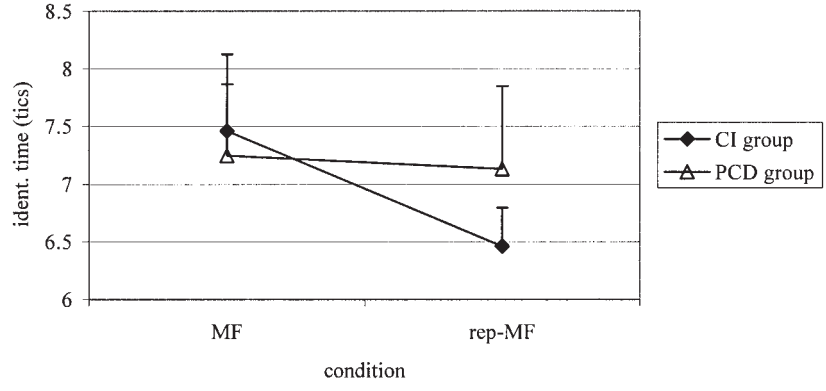


Figure 2. T_1 (i.e., baseline) mean identification (ident.) times (in 16-ms tics) by the cognitively impaired (CI) group versus the preclinically demented (PCD) group on the Perceptual Identification Task: first versus repeated (rep) presentation of the middle-frequency (MF) words. Error bars represent standard error of the mean.

and, third, the Visual Association Test. It may be concluded that dementia was best predicted by little benefit from semantic relations in a cued recall task (and hardly any improvement when words are repeatedly presented), absent implicit remembering of words presented previously, and impaired cued recall of visually interacting objects. Even within a cognitively impaired group of subjects (i.e., a group at-risk for developing dementia, comparable with the MCI concept [most likely the multiple-cognitive-deficits type]), these measures show high accuracy of prediction 2 years before the diagnosis. It may be argued that measures other than the variables commonly used in clinical practice to assess dementia (e.g., MMSE; purely episodic memory tests, such as the 10-Word List-Learning Test or word-recognition tests) were best predictive.

Furthermore, it may be concluded that different measures discriminated best between demented and nondemented subjects 2 years before diagnosis (T_1) compared with at the time of diagnosis (T_2). The only exception was the Visual Association Test, which was a useful measure at both stages. The Word-Recognition Test and the Category Fluency Test differentiated well only at T_2, whereas they showed a relatively low predictive value at T_1. It should be noted that the reduced differentiating value of the Paired-Associate Learning Test at T_2 is most likely explained by the near floor-level performance by the PCD subjects at T_1.

Despite the differentiating value of the Visual Association Test at both stages of the disease (T_1 and T_2), this subtest only seemed useful regarding the assessment of dementia within a cognitively more heterogeneous group. Within the entire group, most nondemented subjects obtained the maximum score on this subtest, whereas the PCD subjects obtained a relatively low score. Within the CI group, relatively low scores were more common. The Visual Association Test seems too rough a measure—mainly because of the low level of difficulty and a limited range of scores—to differentiate well within a more homogeneous group of subjects with cognitive deficits.

The MMSE was found to be less valuable than other measures; it was not among the best discriminating variables in any of the stepwise discriminant analyses. On its own, the MMSE assigned too many subjects incorrectly to the PCD group, even within the large heterogeneous group of elderly subjects. Nonetheless, the MMSE does seem useful as a first and global screening instrument to identify cognitively at-risk subjects (high sensitivity of 78%), for which purpose it was designed. However, analysis of the Visual Association Test showed a similar usefulness as a first dementia screening instrument, with a lower sensitivity of 67% but a higher specificity of 83% (vs. 78% for the MMSE). It should be noted that the Paired-Associate Learning Test on its own better discriminated the PCD subjects from the nondemented subjects (89% sensitivity and 81% specificity). However, it would be less useful as a screening instrument than the Visual Association Test because of its longer administration time (i.e., the Paired-Associate Learning Test consists of 10 items and three trials vs. 6 items and one trial for the Visual Association Test).

The difference in benefiting from semantic relations between PCD subjects and nondemented subjects is an issue frequently investigated in studies testing AD patients (e.g., Hodges & Patterson, 1995; Hodges, Salmon, & Butters, 1990; Rosser & Hodges, 1994; Weingartner, Kawas, Rawlings, & Shapiro, 1993). However, most studies use subjects who are at least mildly demented rather than in a preclinical stage of dementia. Furthermore, these subjects are many years younger and have a higher level of education than those used in the current study. These factors increase the difference between pathological and normal aging processes and, thus, more easily lead to significant differences. Therefore, the strong predictive value of the Paired-Associate Learning Test in the current study—with much smaller differences between the two groups—may be regarded as an important finding for the early assessment of dementia. It may be concluded that poor semantic encoding of to-be-learned information, as has been found in AD patients relative to unimpaired elderly subjects (e.g., Chertkow & Bub, 1990; Monti et al., 1996; Russo & Spinnler, 1994), is also detectable when subjects are still officially nondemented. Sailor et al. (1998) suggested that AD patients have a specific deficit in the ability to evaluate semantic relations. They are no longer able to discriminate between two related concepts, because the attribute knowledge that distinguishes these two concepts has been lost. This explanation may also characterize performance by the PCD subjects in the current study. Some subjects noted that the target word "had something to do with" the cue, but somehow they could not name the correct word. Some subjects named semantically related intrusions, but they repeatedly did not succeed in naming the correct word.

In contrast with the PCD subjects, the nondemented subjects benefited normally from the semantic relation between the words, and they showed a normal learning curve over trials, even if they had cognitive deficits (the CI subjects). The Paired-Associate Learning Test was easier for these unimpaired subjects than was, for example, the 10-Word List-Learning Test, a result that agrees with literature findings—aging effects are greater in free recall tasks than in cued recall or recognition formats (e.g., Bäckman & Wahlin, 1995; Jelicic, Craik, & Moscovitch, 1996; Monti et al., 1996). Thus, unimpaired elderly subjects showed intact performance on cued recall tasks demanding semantic processing. However, they performed deficiently on free recall of semantically unrelated words, which resulted in a relatively small performance difference with PCD subjects. Even when the disease had progressed (at T_2), the performance difference between the two groups was relatively small. Nonetheless, performance on a more passive retrieval task, such as the Word-Recognition Test, showed greater differences and may, therefore, be a more useful diagnostic instrument than an active retrieval task like the 10-Word List-Learning Test. However, the Word-Recognition Test turned out to be "too easy" for PCD subjects (at T_1).

The predictive value of the priming measure of the Perceptual Identification Task seems promising, especially considering the low reliability of this measure (i.e., if the task is improved in this respect, the effect may even be stronger). The value of implicit memory tests may be explained by the limited demands on active, deliberate (conscious) retrieval processes in these tasks. It has frequently been reported that implicit, automatic retrieval processes are intact in normal aging, whereas they have been found to be impaired in AD (see Spaan et al., 2003, for a review). Apparently, this is also true for the preclinical stage of dementia (or AD). However, further research needs to be done considering the findings that AD patients generally perform worse on conceptual priming or *generation* priming tasks rather than perceptual priming or *identification* priming tasks (e.g., Gabrieli et al., 1994; Keane et

al., 1991; Meiran & Jelicic, 1995). Note that the priming measure of the Perceptual Identification Task belongs to the latter category. Therefore, deficits of the PCD subjects might have been more pronounced had a reliable conceptual–generation priming task been available.

Furthermore, it may be concluded that in contrast with episodic, semantic, and implicit memory deficits, short-term (working) memory is intact in PCD. Verbal short-term memory (measured by the Digit Span Task) even showed better performance in PCD subjects than in nondemented subjects (within the entire group of subjects and within the CI group, $p < .05$; see Table 2).

Some final remarks should be made regarding the usefulness of the MCI concept in the early assessment of dementia. As mentioned above, our CI subjects did not officially meet Petersen's MCI criteria (e.g., Petersen et al., 1999) because of absent information regarding the criterion of "subjective memory impairments." We believe it is important to determine the definition of these subjective memory complaints properly. Whether a subject complains depends on the specific question that is asked. We think this criterion of the MCI concept may be interpreted in various ways, depending on the context or perspective that is illustrated in the question(s) asked. This variability influences the nature of the subgroup division and may result in a highly heterogeneous "MCI" subgroup (see also Luis, Loewenstein, Acevedo, Barker, & Duara, 2003, for a critical review). In addition, informant reports (if available) may vary in validity, depending on living status and relationship type toward the subject (e.g., Ready, Ott, & Grace, 2004). Nonetheless, the population-based study of Busse et al. (2003) shows that sensitivity to detect subsequent dementia is increased by excluding the subjective memory complaints criterion, though specificity is reduced.

In addition, it should be emphasized that MCI is usually investigated in elderly persons who visit an outpatient memory clinic rather than in a community-dwelling elderly population, such as the one used in the current study. This may complicate comparisons between studies. Ritchie, Artero, and Touchon (2001) found that MCI was a poor predictor of dementia within a representative population sample. Only 11.1% converted to dementia within a 3-year period, whereas almost all subjects changed category each year. Note that the majority of our subjects (10 of 16) with a CAMDEX minimal dementia classification at T_1 (comparable to the MCI amnestic type [isolated memory deficits]), did not convert to dementia at T_2 (see Footnote 3 and Appendix B, which is available on the Web at http://dx.doi.org/10.1037/0894-4105.19.5.629.supp). In addition, more recent population-based studies of the prevalence of various MCI subtypes show that MCI amnestic type is less frequent than the MCI multiple-cognitive-deficits type (e.g., Busse et al., 2003; Lopez, Jagust, DeKosky, et al., 2003). Lopez, Jagust, DeKosky, et al. (2003) found that 6.0% of their subjects were MCI amnestic, whereas 15.7% had MCI of the multiple-cognitive-deficits type. Furthermore, Busse et al. (2003) found that multiple-cognitive-deficits MCI was the only MCI subtype that showed a significant relative predictive power for onset of dementia 3 years in advance. Thus, limiting MCI criteria to isolated memory impairments seems less efficient in the detection of elderly subjects at risk for developing dementia in the near future.

Naturally, the MCI subjects who actually convert to dementia (or AD) are the most interesting. These subjects are represented in the current study by the PCD subgroup. Thus, cognitive performance of the PCD subjects provides more relevant information on the early predictors of dementia than does performance of an MCI group, in which the identity and the number of subjects that actually develop dementia within a few years is still unknown. Although our CI subgroup (i.e., the cognitively impaired subjects who did not develop dementia 2 years later) did not officially meet MCI criteria (Petersen et al., 1999, 2001), it certainly represented a relevant group that might be encountered in clinical practice because of their deficits. In this way, the CI subjects provided useful reference material in the search for the most sensitive and specific predictors of dementia.

In sum, promising tasks regarding the early assessment of dementia are explicit memory tests requiring semantic processing and implicit memory tests from which repetition priming effects can be derived. Tasks typically used in clinical practice to assess dementia, such as free recall of semantically unrelated items and recognition of words with semantically related distractors, seem less predictive. These tasks may only differentiate between pathological and normal aging when dementia has progressed to a more advanced stage. More specifically, passive episodic retrieval tasks (i.e., recognition) may be useful in an earlier clinical stage than are more active episodic retrieval tasks, including free recall.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bäckman, L., & Wahlin, A. (1995). Influences of item organizability and semantic retrieval cues on word recall in very old age. *Aging and Cognition, 2,* 312–325.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.

Beekman, A. T. F., Deeg, D. J. H., Van Limbeek, J., Braam, A. W., De Vries, M. Z., & Van Tilburg, W. (1997). Criterion validity of the Center for Epidemiologic Studies Depression Scale (CES-D): Results from a community-based study of older subjects in the Netherlands. *Psychological Medicine, 27,* 231–235.

Bird, M., & Luszcz, M. (1991). Encoding specificity, depth of processing, and cued recall in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 13,* 508–520.

Brandt, J., Corwin, J., & Krafft, L. (1992). Is verbal recognition memory really different in Huntington's and Alzheimer's disease? *Journal of Clinical and Experimental Neuropsychology, 14,* 773–784.

Busse, A., Bischkopf, J., Riedel-Heller, S. G., & Angermeyer, M. C. (2003). Subclassifications for mild cognitive impairment: Prevalence and predictive validity. *Psychological Medicine, 33,* 1029–1038.

Carlesimo, G. A., Fadda, L., Lorusso, S., & Caltagirone, C. (1994). Verbal and spatial memory spans in Alzheimer's and multi-infarct dementia. *Acta Neurologica Scandinavica, 89,* 132–138.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type: What do various measures measure? *Brain, 113,* 397–417.

Collie, A., & Maruff, P. (2000). The neuropsychology of preclinical Alzheimer's disease and mild cognitive impairment. *Neuroscience and Biobehavioral Reviews, 24,* 365–374.

Deeg, D. J. H., Beekman, A. T. F., Kriegsman, D. M. W., & Westen-

dorp-de Serière, M. (Eds.). (1998). *Autonomy and well-being in the aging population II: Report from the Longitudinal Aging Study Amsterdam, 1992–1996.* Amsterdam: VU University Press.

Deeg, D. J. H., & Westendorp-de Serière, M. (Eds.) (1994). *Autonomy and well-being in the aging population: Report from the Longitudinal Aging Study Amsterdam, 1992–1993*. Amsterdam: VU University Press.

Deweer, B., Ergis, A. M., Fossati, P., Pillon, B., Boller, F., Agid, Y., & Dubois, B. (1994). Explicit memory, procedural learning and lexical priming in Alzheimer's disease. *Cortex, 30,* 113–126.

Folstein, M. F., Folstein, S. F., & McHugh, P. R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Gabrieli, J. D. E., Keane, M. M., Stanger, B. Z., Kjelgaard, M. M., Corkin, S., & Growdon, J. H. (1994). Dissociations among structural–perceptual, lexical–semantic, and event–fact memory systems in amnesic, Alzheimer's, and normal subjects. *Cortex, 30,* 75–103.

Helkala, E. L., Laulumaa, V., Soininen, H., & Riekkinen, P. J. (1989). Different error pattern of episodic and semantic memory in Alzheimer's disease and Parkinson's disease with dementia. *Neuropsychologia, 27,* 1241–1248.

Hodges, J. R., & Patterson, K. P. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia, 33,* 441–459.

Hodges, J. R., Salmon, D. P., & Butters, N. (1990). Differential impairment of semantic and episodic memory in Alzheimer's and Huntington's diseases: A controlled prospective study. *Journal of Neurology, Neurosurgery, and Psychiatry, 53,* 1089–1095.

Jelicic, M., Craik, F. I. M., & Moscovitch, M. (1996). Effects of ageing on different explicit and implicit memory tasks. *European Journal of Cognitive Psychology, 8,* 225–234.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test* (2nd ed.). Philadelphia: Lea & Febiger.

Keane, M. M., Gabrieli, J. D. E., Fennema, A. C., Growdon, J. H., & Corkin, S. (1991). Evidence for a dissociation between perceptual and conceptual priming in Alzheimer's disease. *Behavioral Neuroscience, 105,* 326–342.

Lindeboom, J., Schmand, B., Tulner, L., Walstra, G., & Jonker, C. (2002). Visual association test to detect early dementia of the Alzheimer type. *Journal of Neurology, Neurosurgery, and Psychiatry, 73,* 126–133.

Lopez, O. L., Jagust, W. J., DeKosky, S. T., Becker, J. T., Fitzpatrick, A., Dulberg, C., et al. (2003). Prevalence and classification of mild cognitive impairment in the Cardiovascular Health Study Cognition Study: Part 1. *Archives of Neurology, 60,* 1385–1389.

Lopez, O. L., Jagust, W. J., Dulberg, C., Becker, J. T., DeKosky, S. T., Fitzpatrick, A., et al. (2003). Risk factors for mild cognitive impairment in the Cardiovascular Health Study Cognition Study: Part 2. *Archives of Neurology, 60,* 1394–1399.

Lowry, R. (2005). Kappa as a measure of concordance in categorial sorting. Retrieved from Vassar College, VassarStats: Web Site for Statistical Computation Web site: http://faculty.vassar.edu/lowry/VassarStats.html

Luis, C. A., Loewenstein, D. A., Acevedo, A., Barker, W. W., & Duara, R. (2003). Mild cognitive impairment: Directions for future research. *Neurology, 61,* 438–444.

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language, 19,* 124–141.

Meiran, N., & Jelicic, M. (1995). Implicit memory in Alzheimer's disease: A meta-analysis. *Neuropsychology, 9,* 291–303.

Monsch, A. U., Bondi, M. W., Butters, N., Paulsen, J. S., Salmon, D. P., Brugger, P., & Swenson, M. R. (1994). A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. *Neuropsychology, 8,* 25–30.

Monti, L. A., Gabrieli, J. D. E., Reminger, S. L., Rinaldi, J. A., Wilson, R. S., & Fleischman, D. A. (1996). Differential effects of aging and Alzheimer's disease on conceptual implicit and explicit memory. *Neuropsychology, 10,* 101–112.

Neri, M., Rubichi, S., DeVreese, L. P., Roth, M., & Cipolli, C. (1998). Validation of the full and short forms of the CAMDEX interview for diagnosing dementia: Evidence from a one-year follow-up study. *Dementia and Geriatric Cognitive Disorders, 9,* 339–346.

Ott, A., Breteler, M. M. B., Van Harskamp, F., Claus, J. J., Van der Cammen, T. J. M., Grobbee, D. E., & Hofman, A. (1995). Prevalence of Alzheimer's disease and vascular dementia: Association with education. The Rotterdam study. *British Medical Journal, 310,* 970–973.

Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., et al. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology, 58,* 1985–1992.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology, 56,* 303–308.

Radloff, L. S. (1977). The CES–D scale: A self-report depression scale developed for research in the general population. *Applied Psychological Measurement, 1,* 385–401.

Randolph, C., Braun, A. R., Goldberg, T. E., & Chase, T. N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's disease: Dissociation of storage and retrieval failures. *Neuropsychology, 7,* 82–88.

Raven, J. C. (1984). *Manual for the Coloured Progressive Matrices* (Rev. ed.). Windsor, England: NFER-Nelson.

Ready, R. E., Ott, B. R., & Grace, J. (2004). Validity of informant reports about AD and MCI patients' memory. *Alzheimer Disease and Associated Disorders, 18,* 11–16.

Rey, A. (1964). *L'examen clinique en psychologie* [Clinical research in psychology]. Paris: Presses Universitaires de France.

Ritchie, K., Artero, S., & Touchon, J. (2001). Classification criteria for mild cognitive impairment: A population-based validation study. *Neurology, 56,* 37–42.

Rosser, A., & Hodges, J. R. (1994). Initial letter and semantic category fluency in Alzheimer's disease, Huntington's disease, and progressive supranuclear palsy. *Journal of Neurology, Neurosurgery, and Psychiatry, 57,* 1389–1394.

Roth, M., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, H., Verma, S., & Goddard, R. (1986). CAMDEX: A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *British Journal of Psychiatry, 149,* 698–709.

Russo, R., & Spinnler, H. (1994). Implicit verbal memory in Alzheimer's disease. *Cortex, 30,* 359–375.

Sailor, K. M., Bramwell, A., & Griesing, T. A. (1998). Evidence for an impaired ability to determine semantic relations in Alzheimer's disease patients. *Neuropsychology, 12,* 555–564.

Savage, R. (1984). *Alphabet Coding Task–15.* Unpublished manuscript, Murdoch University, Perth, Western Australia.

Schacter, D. L. (1992). Understanding implicit memory. *American Psychologist, 47,* 559–569.

Spaan, P. E. J., Raaijmakers, J. G. W., & Jonker, C. (2003). Alzheimer's disease versus normal ageing: A review of the efficiency of clinical and experimental memory measures. *Journal of Clinical and Experimental Neuropsychology, 25,* 216–233.

Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience, 4,* 232–243.

Tombaugh, T. N., & McIntyre, N. J. (1992). The Mini-Mental State Examination: A comprehensive review. *Journal of the American Geriatrics Society, 40,* 922–935.

Trojano, L., Chiacchio, L., De Luca, G., & Grossi, D. (1994). Exploring visuospatial short-term memory defect in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 16,* 911–915.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization and memory* (pp. 381–403). New York: Academic Press.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—Revised.* San Antonio, TX: Psychological Corporation.

Weingartner, H. J., Kawas, C., Rawlings, R., & Shapiro, M. (1993). Changes in semantic memory in early stage Alzheimer's disease patients. *The Gerontologist, 33,* 637–643.

Received March 13, 2003
Revision received December 2, 2004
Accepted December 3, 2004 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write to Demarie Jackson at the address below. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In your letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, "social psychology" is not sufficient—you would need to specify "social cognition" or "attitude change" as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

Write to Demarie Jackson, Journals Office, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242.