



Published in final edited form as:

Alzheimers Dement. 2007 April ; 3(2): 109–121. doi:10.1016/j.jalz.2007.02.002.

The Memory and Aging Telephone Screen (MATS): Development and preliminary validation

Laura A. Rabin, Ph.D.^{a,b}, Andrew J. Saykin, Psy.D.^{a,c,*}, Heather A. Wishart, Ph.D.^a, Katherine E. Nutter-Upham, B.A.^{a,b}, Laura A. Flashman, Ph.D.^a, Nadia Pare, Ph.D.^a, and Robert B. Santulli, M.D.^a

^aNeuropsychology Program, Department of Psychiatry, Dartmouth Medical School/DHMC, Lebanon, NH, USA

^bDepartment of Psychology, Brooklyn College, City University of New York, Brooklyn, NY, USA

^cCenter for Neuroimaging, Department of Radiology, Indiana University School of Medicine, Indianapolis, IN USA

Abstract

Background—Telephone interviews are widely used in geriatric settings to identify eligible research participants and to perform brief follow-up assessments of cognition. This article reports on the development and validation of the Memory and Aging Telephone Screen (MATS), a structured interview for older adults with mild cognitive impairment and/or significant memory complaints. We also developed three alternate forms of the MATS objective memory test to reduce practice effects engendered by multiple administrations.

Methods—Participants were enrolled in a longitudinal study that included 120 older adults with amnesic mild cognitive impairment (MCI), subjective cognitive complaints (CC) but without deficit on neuropsychological tests, and demographically-matched healthy controls (HC). An additional 15 patients with mild probable Alzheimer's disease (AD) completed the alternative forms study. All participants received the original MATS version, and a subset (n = 90) later received two of three alternate forms.

Results—The MATS was sensitive to group differences and the alternate forms were equivalent. MATS objective memory test scores showed adequate stability over one year and were moderately correlated with scores on a widely used list-learning test (CVLT-II).

Conclusions—The MATS, a repeatable telephone screen that includes objective and subjective memory assessments, is useful for detecting individuals in the preclinical and early stages of dementia. Results encourage use of the MATS as a reliable and valid cognitive screening tool in research and clinical settings. Longitudinal assessments are being performed to investigate the predictive validity of the MATS for cognitive progression in MCI.

*Corresponding author. (317) 278-6947, E-mail address: E-mail: asaykin@iupui.edu, Mailing Address: Center for Neuroimaging, Department of Radiology, R2, E124, 950 West Walnut St., Indiana University School of Medicine, Indianapolis, IN USA 46202, Fax: (317) 274-1067.

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

Keywords

telephone screen; mild cognitive impairment; cognitive complaints; neuropsychological assessment; interview; memory

1. Introduction

As the prevalence of dementia increases and strategies for its prevention or delay develop, it is imperative that researchers find cost-effective ways to identify the earliest clinical stages of disease. It is also important to distinguish categories along the continuum from normal aging through varying degrees of cognitive impairment. Brief, efficient screens offer the potential for early detection and discrimination, and can aid in the serial assessment of patients with progressive disorders. When successfully modified for administration over the telephone, screening tools can be used in research settings to identify individuals likely to meet criteria for a particular diagnosis upon more extensive evaluation or to estimate rates of cognitive impairment in a population. These instruments also facilitate frequent and rapid follow-up assessments in longitudinal observational and long-term treatment studies in which many individuals are tracked, sometimes at a distance [1]. Research comparing telephone and face-to-face assessments in older adults consistently has shown that scores on comparable measures may be interchanged [2-5]. Several instruments for screening of cognitive impairment by telephone interview have been validated in older adult populations, providing further support for this strategy.

A common approach is to adapt existing measures of global cognition such as the Mini Mental Status Examination (MMSE) [6], Modified Mini-Mental State Exam (3MS) [7], or Blessed Information-Memory Concentration Test [8] for telephone administration. Research indicates that telephone versions of varying length and complexity provide a reasonable substitute for more costly, in-person evaluations in healthy elders and in those with various forms of dementia or cognitive impairment [9-13]. Additionally, telephone scores can be converted successfully to a Clinical Dementia Rating Scale (CDR) score [14]. A drawback, however, is that global tests may lack diagnostic precision, and they are limited by ceiling effects, particularly in mild and preclinical neurodegenerative disease stages and in high functioning individuals [15-20]. Another approach involves combining several commonly used cognitive tests into a single, telephone-administered battery. Using data from the Religious Orders Study, Wilson and Bennett [4] showed that performance on a battery of seven cognitive tests of episodic, semantic, and working memory was not affected by mode of administration. Knopman and colleagues [21] found that a telephonic cognitive battery consisting of nine brief tests, distinguished, with high precision, healthy elderly from those with cognitive impairment. Other researchers have applied this model successfully in patients with various forms of dementia [22-23]. Overall, these findings suggest that various cognitive measures can be administered reliably to older adults over the telephone. Additionally, in most of the reported research mild hearing impairment did not significantly affect performance. A limitation of the extant literature, however, is that no screening measure has been designed specifically to assess prodromal cognitive deficits or to tap the unique difficulties experienced by individuals in preclinical neurodegenerative disease stages.

Mild cognitive impairment (MCI), often conceptualized as a transitional stage between the cognitive and neurobiological changes associated with normal aging and dementia, is the focus of intense research interest with regard to the early detection of AD and other neurodegenerative processes [24-27]. The prevalence of MCI varies widely from one study to another, and generally ranges from approximately 9 to 18% of individuals older than 65 [28-29]. Although MCI is a heterogeneous condition and may have static or even reversible etiologies, the

memory-predominant subtype, or amnesic MCI (a-MCI) is viewed as an important risk factor for AD. The subtleties in distinguishing MCI from other conditions, such as mild depression and normal aging [30], make it difficult to screen for appropriate candidates for research studies.

Among the most commonly used instruments is the Telephone Interview for Cognitive Status (TICS) [31], which includes tests of orientation, attention, working memory, praxis, sentence repetition, naming to verbal description, recent memory, word opposites, and immediate recall of a 10-word list. Several modified (TICS-m) versions have emerged, which notably include a delayed recall task [32-34], and more recently, two parallel versions of the word list memory task [35]. Although the TICS revision includes items thought to be more sensitive to early cognitive impairment (e.g., delayed recall), it has yet to be utilized extensively in preclinical populations and may have poor positive predictive value in mildly impaired or “ambiguous” cases [34]. Beeri et al. [36] used a Hebrew version of the TICS-m in a sample of Israeli men and found excellent sensitivity for dementia, MCI, and normal cognition (100% for the three comparisons using established cutoff scores), but low specificity (23.5% for each of the conditions). Lines et al. [37] evaluated the predictive utility of the TICS-m for a-MCI, using potential participants from the Prevention of Alzheimer's in Society's Elderly (PRAISE) study who scored within a predefined range. Results indicated that 43% of individuals who passed telephone screening and were seen in the clinic met clinical criteria for a-MCI. Additionally, the delayed recall component of the TICS-m was the most important contributor to the clinical determination of a-MCI, whereas there was a trend toward ceiling effects for the language, attention, and orientation items suggesting that they lack sensitivity. Yaari et al. [38] used data from the Alzheimer's Disease Cooperative Study (ADCS) to investigate recruitment of a-MCI participants for clinical trials. The TICS-m yielded a 51% positive predictive value for participants who met initial cutoff criteria and underwent full clinical evaluation. Further, only five of 26 questions were valuable predictors of a-MCI, and the 10-word delayed recall question alone had a similar predictive accuracy as the entire questionnaire, calling into question the necessity of administering the full TICS-m for recruitment of this clinical population.

We developed a structured telephone screen for use in our longitudinal investigation of cognitive and neural mechanisms in preclinical and early stages of dementia. The Memory and Aging Telephone Screen (MATS) is a 20-minute assessment containing both objective and subjective cognitive variables selected on the basis of conceptual and empirical factors. For example, current diagnostic criteria for a-MCI include a memory complaint and objective evidence of memory impairment in relation to age [39-40]. Therefore, the MATS contains a subjective questionnaire, which inquires about perceived cognitive decline including onset, course, severity relative to peers, and impact on functioning. Additionally, verbal episodic memory tests, particularly list-learning, are considered among the most sensitive measures for early detection of preclinical AD [41-42], whereas other cognitive functions (e.g., language, praxis) may be relatively spared at this stage [24]. Therefore, the MATS includes a verbal memory test with three learning trials, delayed recall, and a recognition condition to afford comprehensive assessment of encoding, storage, and retrieval processes. Because practice effects are problematic in research with repeated assessments meant to document change in cognitive status related to the natural course of a disorder or some treatment or intervention, we created three alternate MATS forms and tested their equivalence. Our overall goal was to determine whether the MATS would confer an advantage over existing telephone screens and facilitate detection of mild memory impairment in our participant groups.

2. Methods

2.1. Participants

Participants were recruited from flyers, newspaper ads, public lectures, and referrals from our medical center clinics. All potential participants who were at least 60 years of age, right-handed, fluent in English, and had a minimum of 12 years of formal education were screened with the MATS and through medical chart review. Exclusion criteria included any significant or uncontrolled medical, psychiatric, or neurological condition (other than MCI) that could affect brain structure or cognition, history of head trauma with loss of consciousness lasting more than five minutes, and current or past history of substance abuse. MATS performance did not affect eligibility. Approximately 40% of screened participants met study inclusion requirements and were enrolled; of those scheduled, over 95% presented for their evaluation. Participants were required to have a knowledgeable collateral informant who could answer questions about their cognition and general health. All participants provided written informed consent according to procedures approved by the Institutional Committee for the Protection of Human Subjects.

Participants underwent detailed neuropsychological evaluation, including measures of memory, attention, executive function, language, spatial ability, general intellectual functioning, and psychomotor speed, as well as standard dementia screens. These included: MMSE; Mattis Dementia Rating Scale, Second Edition (DRS-2) [43]; American National Adult Reading Test (ANART) [44]; Wechsler Adult Intelligence Scale, Third Edition (WAIS-III, Information, Block Design, Digit Span, Digit Symbol, Vocabulary) [45]; Wechsler Memory Scale, Third Edition (WMS-III, Logical Memory, Visual Reproduction) [46]; California Verbal Learning Test, Second Edition (CVLT-II) [47]; Delis Kaplan Executive Function System (DKEFS, Verbal Fluency, Trail Making Test) [48]; Wisconsin Card Sorting Test (WCST, short form) [49]; and the Boston Naming Test (BNT) [50]. Participants also underwent a brief hearing test. All tests were administered by postdoctoral fellows or highly trained psychometric technicians. The mean time between telephone screening with the MATS and in-person evaluation was 8.2 weeks (SD = 8.1).

Level of cognitive complaint was determined from responses on the subjective memory questions from the MATS; Memory Self-Rating Questionnaire [51]; Neurobehavioral Function/Activities of Daily Living Scale (NBF-ADL self- and informant-versions) [52]; Memory Assessment Questionnaire [53-54]; Informant Questionnaire on Cognitive Decline in the Elderly (self- and informant-versions) [55]; and cognitive items from the Geriatric Depression Scale (GDS) [56]. A Cognitive Complaint Index [57] ranging from 0 - 100 was calculated based on the total number of items that could be endorsed across all subjective measures. Participants characterized as having significant cognitive complaints typically endorsed 20% or more of the index items. A Board-certified geropsychiatrist conducted a semi-structured interview to rule out depression or other psychiatric disorders. Participants underwent structural brain MRIs, which were reviewed by a Board-certified neuroradiologist to rule out incidental pathology that could account for cognitive symptoms.

A panel of neuropsychologists and a geropsychiatrist reviewed the evaluation results at a weekly consensus conference to determine group classification according to the criteria outlined in Box 1. Though MATS scores were available and considered during the consensus conference, diagnostic decisions were based on performance on standardized tests of episodic memory and other cognitive abilities, and not on MATS results. In addition to healthy controls (HC) and amnesic mild cognitive impairment (MCI), we also included a group of non-depressed elders who present with significant self-reported and/or informant-reported cognitive complaints (CC), but who perform normally on neuropsychological testing. Recent research suggests that CCs show structural brain changes intermediate between those seen in

MCI and healthy older adults without such complaints, and therefore may represent a pre-MCI population worthy of close investigation and longitudinal follow [57].

Study participants were consecutively enrolled and included 39 patients with a-MCI,¹ 38 CCs, and 43 demographically-matched HCs. Sixty-six percent of participants were recruited from flyers, public lectures, and newspaper advertisements, 20% were referred by physicians, and 14% learned about the study from friends or other participants. Notably, a disproportionate amount of MCI participants (38.5%) were referred by physicians as compared to CC and HCs (13.2% and 9.3%, respectively) [$\chi^2(4, N = 120) = 15.04, p = .005$]. The primary factor leading most participants to volunteer was concern about their memory (67.5%), followed by an interest in furthering research (28%), and “other” reasons (< 5%). MCI and CC participants disproportionately endorsed “concern about memory” (84% and 82%, respectively) relative to HCs (40%) [$\chi^2(4, N = 120) = 24.05, p < .001$]. Sixty six percent of study informants were identified as participants' spouses or significant others, 17.5% were participants' sons or daughters, and the remaining 16.5% were identified as friends or other family members.²

2.2. Measure

The Memory and Aging Telephone Screen (MATS), outlined in Boxes 2 and 3, is divided into six sections: (1) basic demographic information, (2) current and past medical and psychiatric history, (3) pre-MRI metal screen, (4) a description of our Memory and Aging study, (5) a 12-item, forced-choice, subjective memory questionnaire, and (6) a 10-item list learning test with three learning/immediate recall trials, a delayed recall trial, and a delayed yes/no recognition condition. While designed as a telephone instrument, the MATS can be used in face-to-face contexts or for screening the visually impaired or those unable to read or write. With the exception of the objective memory test, all MATS items can be administered to a reliable collateral informant or caregiver. Participants should be able to hear spoken language at a conversational volume and comprehend and follow basic instructions. Examiners should speak clearly and articulate distinctly, and participants should be instructed to complete the screen in a quiet, distraction-free environment. Administration instructions are printed on the MATS protocol, which takes approximately 20 minutes to complete. Box 3 presents a summary of the 12-item subjective memory questionnaire including scoring procedures. All questions are read aloud, and there is space on the protocol to record additional information spontaneously offered by participants.

For the objective memory test participants listen to a list of words, read at a rate of approximately one word per second, and repeat back as many words as possible in any order (learning trial 1). For learning trials 2-3, participants again listen to the words and repeat back as many as possible, in any order, including words already said. Participants are not explicitly informed about the delayed memory tasks. After an approximately 12-minute interval consisting of questions about general health and family medical history, participants recall as many of the words as they can (i.e., delayed recall). Participants then complete a recognition trial, which consists of 10 target items randomly mixed with 30 distractors; participants respond “yes” to words from target list and “no” to words not from that list. Examiners record responses verbatim and may derive separate scores for the following variables: three learning/immediate recall trials (correct / 10); total learning score for trials 1 - 3 (correct / 30); delayed recall (correct / 10); repetition errors for trials 1 - 3 (total errors) and delayed recall (total errors); and intrusion errors for trials 1 - 3 (total errors) and delayed recall (total errors). For the delayed recognition condition, scores include the number of true positive, false positive, true negative, and false negative responses; a discrimination index also is calculated as $[TP - (.33 * FP)]$.

¹Please note that participants in the current paper were diagnosed with a-MCI, and all subsequent references to MCI relate to this subtype (except where otherwise indicated).

²Refer to Saykin et al. [45] for further details regarding this cohort (overlap > 85% with participants in the current study).

The initial MATS objective memory test consisted of ten, semantically unrelated, one- and two-syllable nouns (five of each), which were easily pronounced and perceived over the telephone. We later developed three alternate test forms using words that approximated the original list on number of syllables, part of speech (all words were concrete nouns), number of letters, number of phonemes, and frequency of use (derived from Francis and Kucera [58]). Box 4 presents the stimuli from the learning trials of the four parallel forms, and Table 1 shows equivalency data related to the primary matching variables. Attention also was paid to secondary matching variables including item familiarity, imageability, and meaningfulness as provided in the MRC Psycholinguistic database [http://www.psych.rl.ac.uk/MRC_Psych_Db_files/psych.htm]. There were no statistically significant differences among the four test forms on any of the primary or secondary matching variables ($p > .05$). As with the original list, an important consideration was to select words that were easy to distinguish over the telephone. After preliminary testing with a random sample of seven study participants, we discarded 15 words that were difficult for participants to understand or for examiners to pronounce.

2.3. Overview of analyses

In the first part of the study (*Validation*), we examined group differences in MATS scores on the subjective and objective memory assessments. For the MATS objective memory test, we further investigated correlations with another widely used test of verbal episodic memory (i.e., CVLT-II), and calculated test-retest stability over a 12-month period. In the second part of the study (*Alternate test forms*) we developed three alternate forms of the MATS objective memory test to reduce task-specific practice effects caused by multiple administrations [59-60]. After construction of the alternate MATS forms we investigated their equivalence and ability to detect group differences in a subset ($N = 90$) of our participants.

3. Results

3.1 Validation study

Statistical analyses were performed using SPSS Version 12.0.1 for Windows (SPSS Inc., Chicago, IL). Select demographic variables are presented in Table 2. There were no significant group differences in age, education, gender, or estimated baseline Verbal IQ (i.e., ANART score). As expected based on study classification criteria, scores on the DRS-2 and MMSE were reduced in MCI relative to the other groups, though MCI participants scored above the cutoff for dementia. Similarly, MCI participants showed significant deficits on tests of memory (e.g., DRS-2, CVLT-II, WMS-III) relative to HCs. The CC group showed an intermediate pattern of performance on memory tests, with scores typically falling within normal limits (i.e., within 1SD of HC). Though no participant was clinically depressed or scored in the depressed range on the GDS, CC and MCI participants tended to endorse several more items than HCs.

Degrees of asymmetry and peakedness in the distribution of MATS scores were evaluated by calculating skew and kurtosis statistics, applying cutoff values recommended by Tabachnick and Fidell [61]. All MATS subjective and objective memory test scores were normally distributed with the exception of delayed recognition (i.e., discrimination), which was negatively skewed (-.95) indicating a mild ceiling effect. Parametric tests were conducted throughout using two-tailed tests, and nonparametric equivalents also were utilized in analyses involving discrimination scores.

ANOVA was used to evaluate group differences in MATS scores on the subjective memory test. Scores were derived by summing 10 of the 12 items³ (range = 0 to 10 points). Results were statistically significant $F(2, 117) = 28.72, p < .001$, revealing the overall effect of diagnostic group membership (see Table 3 and Figure 1). *Post-hoc* comparisons using the

Tukey HSD correction indicated that HCs endorsed fewer subjective memory complaints than CC and MCI participants, who did not differ from each other. One-way analysis of covariance (ANCOVA) also was used to evaluate mean differences, with the same overall pattern of findings $F(2, 113) = 8.62, p < .001$. Results indicated that gender, education, and depressive symptoms (i.e., adjusted GDS score) did not account for a significant amount of the variance between groups ($p > .05$), while age was a significant covariate ($p < .01$), but only accounted for approximately 7% of the variance ($\eta^2 = .069$).

MANOVA was used to evaluate group differences in MATS scores on the objective memory test including total recall across the three learning trials (range = 0 to 30 points), total recall on the delay condition (range = 0 to 10 points), and discrimination on the delayed recognition test (range = 0 to 10 points). The overall multivariate test (Wilks's Lambda) was significant, $F(2, 117) = 20.25, p < .001$, and omnibus univariate tests were significant for all three variables ($p < .001$) (see Table 3). *Post-hoc* comparisons using the Tukey HSD correction indicated that for the learning and delay recall trials, HCs obtained higher scores than CCs, who obtained higher scores than MCI participants. For the recognition condition, HC and CC participants both achieved better discrimination than MCIs. Table 3 presents means and standard deviations by diagnostic group along with significance tests and effect sizes for each MATS variable analyzed.⁴ Multivariate analysis of covariance (MANCOVA) also was used to evaluate mean differences, with the same overall pattern of findings, $F(2, 113) = 8.62, p < .001$. Results indicated that gender, education, and adjusted GDS score did not account for a significant amount of the variance between groups ($p > .05$), while age was a significant covariate ($p < .05$), but only accounted for approximately 7% of the variance. Further analyses using Pearson correlational coefficients indicated that of the three objective memory test variables, only Total Learning showed a statistically significant relationship with age ($r = -.29, p < .01$). Age was not significantly related to Total Score on the subjective memory questionnaire ($r = .06, p > .05$).

We examined relations between scores on the MATS objective memory test and the CVLT-II, a widely used measure of verbal episodic memory with a similar list-learning format. Pearson correlation coefficients were moderate for total number of words learned across trials ($r = .60, p < .001$), delayed recall ($r = .60, p < .001$), and delayed recognition ($r = .47, p < .001$). Stability of performance on the MATS objective memory test was assessed in 36 of the 43 HC participants with available re-test data over a 12-month period. Stability coefficients were high for total learning score ($r = .70, p < .001$), and this was comparable with test-retest correlations reported for learning scores on the Hopkins Verbal Learning Test ($r = .50$) [62] and CVLT-II ($r = .82$) [47]. Stability coefficients were moderate for the delayed recall ($r = .48, p = .003$) and discrimination ($r = .41, p = .013$) conditions. A two-tailed, paired t-test indicated that the average mean difference in total learning between Time 1 and Time 2 ($-0.61, SD = 2.72$) was not statistically significant $t(35) = -1.35, p = .19$. Average mean differences for the delay ($-0.14, SD = 1.82$) and delayed recognition ($-0.15, SD = 1.09$) conditions also were not statistically significant, $t(35) = -0.46, p = .650, t(35) = -0.81, p = .42$.

3.2. Alternate test forms study

To select participants for the alternate test form study, we used a stratified random sample and supplemented refusals by group membership. Seventy-five participants from the validation study received two alternate forms of the MATS objective memory test on separate occasions,

³Two of the twelve questions (i.e., items 5 and 6) were recently added to the test protocol and therefore were not included in the current analysis due to the large amount of missing data.

⁴Due to the high degree of skewness in the recognition discriminability data, a Kruskal-Wallis test was conducted, with significant results: $\chi^2(2, N = 120) = 33.18, p < .001$. Pairwise comparisons using Mann-U Whitney tests indicated the following pattern of results: HC > CC > MCI.

approximately one day apart (mean interval between administrations = 27.3 hours, SD = 9.6). We also included an additional 15 Memory and Aging Study participants who met criteria for probable mild AD, as defined by the NINCDS-ADRDA criteria [63]. Nine randomly selected participants were either unreachable or unwilling to take part in this process, and other participants were randomly selected to take their place. The final sample consisted of 15 AD, 25 a-MCI, 25 CC, and 25 HC participants. The AD group was demographically equivalent to the other three groups ($p > .05$) with regard to age ($M = 75.07$, $SD = 7.52$), education ($M = 15.47$ yrs, $SD = 3.20$), and gender distribution (Males = 7, Females = 8). As expected based on diagnostic criteria, the AD group attained significantly lower scores on tests of general cognitive functioning (MMSE / 30: $M = 22.52$, $SD = 4.50$) and episodic verbal memory (CVLT-II Learning / 80: $M = 23.27$, $SD = 6.53$) as compared to HCs, CCs, and MCIs.

Because the alternate test forms were administered during longitudinal data collection, participants were at various stages of involvement in the Memory and Aging Study including: *Year 1*: 37, *Year 2*: 27, *Year 3*: 19, and *Year 4*: 6, and *Year 5*: 1 participant. The number of participants in each year of the study did not differ across participant groups $\chi^2(12, N = 90) = 13.06$, $p = .37$. Participants received two of three newly developed test forms in addition to Form 1, which they had completed upon initial screen or during longitudinal follow (within 8 weeks of completing Forms 2-4). Thus, Form 1 was completed by all 90 participants, while Forms 2, 3, and 4 were completed by 60 participants each, with a total of 30 participants completing each combination of test versions: Forms 1, 2, 3; Forms 1, 3, 4, or Forms 1, 2, 4.

One-way ANOVAs and repeated measures one-way ANOVAs were used to investigate inter-form equivalence among the four versions of the MATS objective memory test. As shown in Table 5 and Figure 1, the forms did not differ significantly with respect to the learning, delayed recall, or recognition conditions ($p > .05$). MANOVA also was used to evaluate group differences in objective memory test performance (on Form 1) in this subset of study participants. The overall multivariate test (Wilks's Lambda) was significant, $F(3, 86) = 11.98$, $p < .001$, and omnibus univariate tests were significant for all learning, delayed recall, and delayed recognition variables ($p < .001$). *Post-hoc* comparisons using the Tukey HSD correction indicated that for the learning and delayed recognition conditions, HC and CC participants achieved higher scores than MCI, who in turn achieved higher scores than AD patients. For the delayed recall condition, HC and CC participants scored higher than MCI and AD, who did not differ from each other.⁵ Table 4 presents means and standard deviations by diagnostic group along with significance values and posthoc test results. Group differences in objective memory test performance also were investigated for Forms 2 - 4, though the number of participants per group was much smaller for these analyses (N ranged from 16 to 18 for HC, CC, and MCI groups; N = 10 for AD group). The same pattern of mean differences held up across the three alternate test forms for the HC, CC, and AD groups (HC, CC > AD) on learning, delayed recall, and delayed recognition conditions; MCI participants showed an inconsistent pattern of scores, and generally performed better than they did on Form 1 (at the level of CC or HCs).

4. Discussion

While several telephonic screening tools have been validated in dementia populations, the literature is limited on their utility in preclinical disease stages. We reported on the development of the Memory Assessment Telephone Screen (MATS), an instrument designed to screen individuals with mild cognitive impairment and/or significant complaints about cognition. The

⁵Due to the high degree of skewness in the recognition discriminability data, a Kruskal-Wallis test was conducted, with significant results: $\chi^2(3, N = 90) = 47.54$, $p < .001$. Pairwise comparisons using Mann-U Whitney tests indicated the following pattern of results: HC, CC > MCI > AD.

MATS contains a 10 item list-learning test with a delayed recall condition-- the variable previously shown to have the strongest predictive accuracy for amnesic MCI [38]. The encoding phase of the list-learning test consists of three learning trials (instead of one trial) to maximize the potential for elaborate encoding and to reduce the likelihood that brief attentional lapses will contribute to poor performance. We also include a forced-choice recognition condition to enable further examination of retrieval (as opposed to encoding) deficits and evaluation of response bias.

The MATS was not modeled after the MMSE and excludes items known to lack sensitivity in preclinical groups (e.g., orientation to person or place, basic expressive language, praxis). In addition to the objective memory test, the MATS contains a forced-choice subjective memory questionnaire to reflect current conceptualizations of MCI, which tend to place greater emphasis on memory complaints and perceived decline in memory ability [26,64]. In addition, researchers have begun to study a group of individuals with a history of cognitive complaints (CC) and mildly declining function, who do not yet perform 1.5 SD below the mean of healthy controls on psychometric memory tests. Research from our lab and others [57,65] indicates these individuals show structural and functional brain changes intermediate between MCI and HCs, and may be classified as "pre MCI." Therefore, inclusion of both objective and subjective memory assessments may confer an advantage over existing screens in terms of the ability to detect and characterize a wider spectrum of older adults presenting with significant cognitive complaints and/or actual performance deficits.

The MATS was well-tolerated in our group of 135 nondepressed older adults, and did not show a ceiling effect (even in cognitively intact controls). Moderate stability of learning, delayed recall, and delayed recognition conditions was observed for Form 1 over a 12-month period in HCs, suggesting that the MATS is not particularly susceptible to explicit learning effects. The moderate associations with various CVLT variables provided preliminary evidence of concurrent validity. With the exception of Total Learning score, the MATS did not show a significant relationship with age on either the objective or subjective variables. Similar findings were reported for the 12-item HVLT and may reflect the relative ease of the MATS as compared to the lengthier CVLT, on which age is correlated with both learning and delayed recall trials [66]. Another possibility is that the relatively high education level of our participants obscured expected age-related declines often seen on episodic memory tests. While education, gender, and depressive symptoms did not significantly influence current results, future research will examine relations between the MATS and these variables in a more diverse sample.

An important study goal was to investigate the MATS' ability to detect differences in older adults with varying degrees of cognitive compromise. The MATS objective memory test was sensitive to group differences, with MCI participants scoring approximately 1.5 SDs below the mean of HCs across memory test conditions (clinically impaired range). The CC group showed an intermediate level of performance, generally scoring about .7 SDs below the mean of HCs, which is considered within normal limits clinically (low average to average range). When 15 mild AD patients were included in the alternate forms phase of the study, the pattern of group differences was as follows: HC, CC > MCI > AD, with ADs generally scoring more than 2 SDs below the mean of HCs (clinically impaired range). On the subjective memory questionnaire, CC and MCI participants endorsed approximately twice as many complaints as HCs, as expected based on study criteria. Overall, findings supported the validity of the MATS as an indicator of cognitive impairment in our clinical groups. Longitudinal follow-up of MCI and CCs also will be necessary to confirm the MATS' diagnostic value and to monitor rates of progression from CC to MCI or MCI to AD, in order to determine which variables best predict clinical conversion. It is important to note that the MATS would likely be less sensitive to early cognitive impairment that presents primarily in the form of nonamnesic deficits (e.g., visuospatial dysfunction, constructional apraxias). Future research should examine the MATS

in other MCI subtypes (e.g., multiple domain) [67] or patient populations (e.g., vascular dementia, delirium), and modify test items as required. Notably, the MATS has been utilized successfully with patients in more advanced disease stages (i.e., various forms of dementia), and adapted for our other clinical samples such as traumatic brain injury and breast cancer.

In order to avoid practice-related measurement error of participants undergoing serial testing as part of our longitudinal study, we developed three alternate forms of the MATS objective memory test (yielding a total of four forms). The words selected matched the original list on various criteria (e.g., syllables, frequency of usage), and were easy to pronounce and distinguish over the telephone. Analyses of the learning, delayed recall, and recognition conditions indicated that all four forms were interchangeable, confirming their utility in research requiring repeated testing of verbal episodic memory. Additionally, group differences generally held across the four versions, with the exception of MCI who tended to show improved performance on the alternate test forms. Future research should employ a larger sample of MCI patients and apply a between-subjects design to ascertain whether the higher scores reflect a true form difference or are attributable to other factors (e.g., artifact of small sample size, differential benefit in MCI from the short test-retest interval related to anticipatory effects or use of external aids). Further, the test-retest interval for the two alternate forms was shorter than the interval between the original and alternate forms. While this was unavoidable in our current longitudinal study, future research should attain a more consistent time interval between all test versions.

The MATS shares limitations with some existing telephone instruments in that level of motivation, listening and privacy conditions, and auditory deficits may affect performance. While telephone interviewers should acquire an ability to attend to issues like disinterest, distractibility, or cheating (e.g., writing down memory test items, receiving assistance), in our experience, these potential problems are detectable and manageable. We did not directly compare the MATS with existing telephone instruments such as the TICS-m and therefore cannot determine if the MATS is more useful in actual practice. Additionally, we applied strict entry criteria and our sample was predominantly Caucasian, limiting the generalizability of our results. Use of the MATS in more diverse groups may change its validity, and future research should determine the degree to which medical and demographic variables affect response levels; appropriate adjustments or norms for specific groups may be required. Administration time for the MATS is approximately 20 minutes, and may be too lengthy for certain settings. If some of the information-gathering sections are omitted, the time required for the memory assessments (including the delay) could be reduced to approximately 15 minutes. Based on our experience, however, the history sections are useful for gathering relevant patient information and for establishing rapport. Finally, screening instruments help identify individuals with a high probability of having a problem who would benefit from in-person evaluation; they are not meant to provide a definitive diagnosis or substitute for comprehensive assessment.

Some additional directions for research warrant mention. Alternative communication modes such as videoconferencing and computer-automated or computer-assisted telephone screening, while at an early stage of development, show promise for the future [68-70]. Although it might be possible to adapt the MATS to one of these formats, it remains to be seen whether these technologies become widely available, cost-effective, and readily embraced by older adults. It is also important to note that the MATS may not be appropriate for all individuals. For example, in those with significant auditory difficulties, hearing amplification devices may be required for successful administration. Individuals with more advanced cognitive decline may require informant rather than patient versions of telephone screens. Mintzer et al. [71] found that a dementia screening tool based on caregiver responses showed high reliability with subsequent clinical evaluation in a small community-based sample. To this end, all MATS items, with the exception of the objective memory test, can be administered using a reliable informant. While

we have analyzed key variables from the MATS, consideration may be given to additional information such as primacy and recency effects, direct comparison of recall versus recognition scores, and responses to specific subjective memory questions. Incorporation of this type of information into the scoring system may allow for the identification of more discrete cognitive constructs. Finally, we hope to determine the classification accuracy of the MATS and develop cutoff scores with high sensitivity and specificity. Our preliminary findings suggest that a cutoff score of ≤ 33 of 50 points on the objective memory yields the best sensitivity/specificity for MCI. This value, however, will likely need to be adjusted in other settings, in research applying different diagnostic criteria for MCI (e.g., the European Consortium on Alzheimer's Disease [64] or Alzheimer's Disease Neuroimaging Initiative [72]), or in studies of frank dementia.

Summary

Telephone testing provides a practical and efficient mechanism for screening of cognitive impairment in elderly populations, where factors such as physical impairment, financial limitations, or geographic dispersion may affect the feasibility of in-person contact during recruitment or longitudinal follow. The MATS, with its ease of administration and scoring, tolerability, reliability, and discriminative validity can help identify individuals in the preclinical or mild stages of dementia, when treatment that prevents or slows cognitive decline can exert the strongest impact. In addition to the subjective and objective memory assessments, the MATS gathers qualitative information about a variety of medical, psychological, and social issues that may be useful in clinical or research settings, and which can be administered telephonically or in-person by physicians, psychologists, nurses, or other trained personnel. Moreover, the existence of four comparable MATS forms enables serial testing. While initial cross sectional results are encouraging, only upon subsequent follow-up will it be possible to determine the predictive validity of the MATS for cognitive progression in our clinical groups.

Acknowledgments

The authors are grateful to the following individuals for their help with this study: Leslie Baxter, Cheryl Brown, Margaret Nordstrom, Heather Pixley, and Paul Wang of the Department of Psychiatry. Supported, in part, by grants from the National Institute on Aging (R01 AG19771), Alzheimer's Association (IIRG-94-133; IIRG-99-1653, sponsored by the Hedco Foundation), the Hitchcock Foundation, the Ira DeCamp Foundation, and New Hampshire Hospital, Concord, NH.

References

1. Lipton RB, Katz MJ, Kulansky G, Sliwinski MJ, Stewart WF, Verghese J, et al. Screening for dementia by telephone using the Memory Impairment Screen. *J Am Geriatr Soc* 2003;51:1382–1390. [PubMed: 14511157]
2. Kawas C, Karagiozis BA, Resau L, Corrada M, Brookmeyer R. Reliability of the Blessed Telephone Information-Memory-Concentration Test. *J Geriatr Psychiatry Neurol* 1995;8:238–242. [PubMed: 8561839]
3. Roccaforte WH, Burke WJ, Bayer BL, Wengel SP. Reliability and validity of the Short Portable Mental Status Questionnaire administered by telephone. *J Geriatr Psychiatry Neurol* 1994;7:33–38. [PubMed: 8192828]
4. Wilson RS, Bennett DA. Assessment of cognitive decline in old age with brief tests amenable to telephone administration. *Neuroepidemiology* 2005;25:19–25. [PubMed: 15855801]
5. Zhou J, Zhang X, Mundt JC, Wang L, Meng C, Chu C, et al. A comparison of three dementia screening instruments administered by telephone in China. *Dementia* 2004;3:69–81.
6. Folstein MF, Folstein SE, McHugh PR. Mini-Mental state: A practical method for grading the cognitive state of patients for the clinicians. *J Psychiatr Res* 1975;12:189–198. [PubMed: 1202204]
7. Teng EL, Chui HC. The modified mini mental state (3MS) examination. *J Clinical Psychiatry* 1987;48:314–318. [PubMed: 3611032]

8. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral gray matter of elderly subjects. *Br J Psychiatry* 1968;114:797–811. [PubMed: 5662937]
9. Norton MC, Tschanz JA, Fan X, Plassman BL, Welsh-Bohmer KA, West N, et al. Telephone adaptation of the Modified Mini-Mental State Exam (3MS). The Cache County Study Neuropsychiatry *Neuropsychol Behav Neurol* 1999;12:270–276.
10. Lanska DJ, Schmitt FA, Stewart JM, Howe JN. Telephone-assessed mental state. *Dementia* 1993;4:117–119. [PubMed: 8358512]
11. Newkirk LA, Kim JM, Thompson JM, Tinklenberg JR, Yesavage JA, Taylor JL. Validation of a 26-point telephone version of the Mini-Mental State Examination. *J Geriatr Psychiatry Neurol* 2004;17:81–87. [PubMed: 15157348]
12. Mettieri T, Geroldi C, Pezzini A, Frisoni GB, Bianchetti A, Trabucchi M. The ITEL-MMSE: An Italian telephone version of the Mini-Mental State Examination. *Int J Geriatr Psychiatry* 2001;16:166–167. [PubMed: 11241721]
13. Roccaforte WH, Burke WJ, Bayer BL, Wengel SP. Validation of a telephone version of the Mini-Mental State Examination. *J Am Geriatr Soc* 1992;40:697–702. [PubMed: 1607586]
14. Go RCP, Duke LW, Harrell LE, Cody H, Bassett SS, Folstein MF, et al. Development and validation of a Structured Telephone Interview for Dementia Assessment (STIDA): The NIMH Genetics Initiative. *J Geriatr Psychiatry Neuro* 1997;10:161–167.
15. De Jager C, Milwain E, Budge M. Early detection of isolated memory deficits in the elderly: The need for more sensitive neuropsychological tests. *Psychol Med* 2002;32:483–491. [PubMed: 11989993]
16. Hayes R, Gordon B, Selnes O, Hasenauer D. Lack of sensitivity of the Mini-Mental State Exam for detecting cognitive impairments. *Neurology* 1992;40(Suppl 3):220.
17. Tang-Wai DF, Knopman DS, Geda YE, Edland SD, Smith GE, Ivnik RJ, et al. Comparison of the short test of mental status and the mini-mental state examination in mild cognitive impairment. *Arch Neurol* 2003;60:1777–1781. [PubMed: 14676056]
18. Tierney MC, Szalai JP, Dunn E, Geslani D, McDowell I. Prediction of probable Alzheimer disease in patients with symptoms suggestive of memory impairment. Value of the Mini-Mental State Examination *Arch Fam Med* 2000;9:527–532.
19. Tombaugh TN, McIntyre NJ. The Mini-Mental State Examination: A comprehensive review. *J Am Geriatr Soc* 1992;40:922–935. [PubMed: 1512391]
20. White N, Scott A, Woods RT, Wenger GC, Keady JD, Devakumar M. The limited utility of the Mini-Mental State Examination in screening people over the age of 75 years for dementia in primary care. *Br J Gen Pract* 2002;52:1002–1003. [PubMed: 12528587]
21. Knopman DS, Knudson D, Yoes ME, Weiss DJ. Development and standardization of a new telephonic cognitive screening test: The Minnesota Cognitive Acuity Screen (MCAS). *Neuropsychiatry Neuropsychol Behav Neurol* 2000;13:286–296. [PubMed: 11186165]
22. Monteiro IM, Boksay I, Auer SR, Torossian C, Sinaiko E, Reisberg B. Reliability of routine clinical instruments for the assessment of Alzheimer's disease administered by telephone. *J Geriatr Psychiatry Neurol* 1988;11:18–24. [PubMed: 9686748]
23. Gatz M, Reynolds CA, John R, Johansson B, Mortimer JA, Pedersen NL. Telephone screening to identify potential dementia cases in a population-based sample of older adults. *Int Psychogeriatr* 2002;14:273–289. [PubMed: 12475088]
24. Arnaiz E, Almkvist O. Neuropsychological features of mild cognitive impairment and preclinical Alzheimer's disease. *Acta Neurol Scand* 2003;107(Suppl 179):34–41.
25. Blennow K. CSF biomarkers for mild cognitive impairment. *J Intern Med* 2004;256:224–234. [PubMed: 15324365]
26. Winblad B, Palmer K, Kivipelto M, Jelic V, Fratiglioni L, Wahlund LO, et al. Mild cognitive impairment – beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. *J Intern Med* 2004;256:240–246. [PubMed: 15324367]
27. Wolf H, Jelic V, Gertz HJ, Nordberg A, Julin P, Wahlund LO. A critical discussion of the role of neuroimaging in mild cognitive impairment. *Acta Neurol Scand* 2003;107(Suppl 179):52–76.

28. DeCarli C. Mild cognitive impairment: Prevalence, prognosis, aetiology, and treatment. *Lancet Neurol* 2003;2:15–21. [PubMed: 12849297]
29. Smith, G.; Rush, BK. Normal aging and mild cognitive impairment. In: Attix, DK.; Welsh-Bohmer, KA., editors. *Geriatric neuropsychology: Assessment and intervention*. New York: The Guilford Press; 2006. p. 27-55.
30. Jelic V, Kivipelto M, Winblad B. Clinical trials in mild cognitive impairment: lessons for the future. *J Neurol Neurosurg Psychiatry* 2006;77:429–438. [PubMed: 16306154]
31. Brandt J, Spencer M, Folstein M. The Telephone Interview for Cognitive Status. *Neuropsychiatry Neuropsychol Behav Neurol* 1988;1:111–117.
32. Breitner JC, Welsh KA, Magruder KM, Churchill CM, Robinette CD, Folstein MF. Alzheimer's disease in the National Academy of Sciences Registry of aging twin veterans: Pilot Investigations. *Dementia* 1990;1:297–303.
33. De Jager CA, Budge MM, Clarke R. Utility of TICS-M for the assessment of cognitive function in older adults. *Int J Geriatr Psychiatry* 2003;18:318–324. [PubMed: 12673608]
34. Welsh KA, Breitner JC, Magruder-Habib KM. Detection of dementia in the elderly using telephone screening of cognitive status. *Neuropsychiatry Neuropsychol Behav Neurol* 1993;6:103–110.
35. Hogervost E, Bandelow S, Hart J, Henderson VW. Telephone word-list recall tested in the Rural Aging and Memory study: two parallel versions for the TICS-M. *Int J Geriatr Psychiatry* 2004;19:875–880. [PubMed: 15352146]
36. Beerli MS, Werner P, Davidson M, Schmidler J, Silverman J. Validation of the modified Telephone Interview for Cognitive Status (TICS-m) in Hebrew. *Int J Geriatr Psychiatry* 2003;18:381–386. [PubMed: 12766912]
37. Lines CR, McCarroll KA, Lipton RB, Block GA. Telephone screening for amnesic mild cognitive impairment. *Neurology* 2003;60:261–266. [PubMed: 12552041]
38. Yaari R, Fleisher AS, Gamst AC, Bagwell VP, Thal LJ. Utility of the telephone interview for cognitive status for enrollment in clinical trials. *Alzheimer's & Dementia* 2006;2:104–109.
39. Petersen RC, Doody R, Kurz A, Mohs R, Morris JC, Rabins PV, et al. Current concepts in mild cognitive impairment. *Arch Neurol* 2001;58:985–1992.
40. Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL, DeKosky ST, et al. Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review): Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology* 2001;56:1133–1142. [PubMed: 11342677]
41. De Jager CA, Hogervorst E, Combrinck M, Budge MM. Sensitivity and specificity of neuropsychological tests for mild cognitive impairment, vascular cognitive impairment, and Alzheimer's disease. *Psychol Med* 2003;33:1039–1050. [PubMed: 12946088]
42. Pare, N.; Saykin, AJ.; Rabin, LA.; Wishart, HA.; Flashman, LA.; Nutter-Upham, KE., et al. Differential sensitivity and specificity of memory measures for classification of MCI. Proceedings of the 34th Annual Meeting of the International Neuropsychological Society; 2006. p. 55
43. Jurica, P.; Leitten, C.; Mattis, S. *Dementia Rating Scale-2*. Lutz, FL: Psychological Assessment Resources; 2001.
44. Grober E, Sliwinski M. Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *J Clin Exp Neuropsychol* 1991;13:933–949. [PubMed: 1779032]
45. Wechsler, D. *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: The Psychological Corp; 1997.
46. Psychological Corporation. *WMS-III technical manual*. San Antonia, TX: Author; 1997. Wechsler Memory Scale-Third Edition.
47. Delis, DC.; Kramer, JH.; Kaplan, E.; Ober, BA. *The California Verbal Learning Test-Second Edition: Adult version manual*. San Antonio, TX: Psychological Corporation; 2000.
48. Delis, D.; Kaplan, E. *Delis-Kaplan Executive Function Battery*. San Antonio, TX: Psychological Corporation; 2001.
49. Heaton, RK.; Chelune, GJ.; Talley, JL.; Kay, GG.; Curtiss, G. *Wisconsin Card Sorting Test manual: Revised and expanded*. Lutz, FL: Psychological Assessment Resources; 1993.

50. Goodglass, H.; Kaplan, E.; Barresi, B. Boston Diagnostic Aphasia Examination-Third Edition. Philadelphia: Lippincott Williams & Wilkins; 2001.
51. Squire LR, Wetzel CD, Slater PC. Memory complaints after electroconvulsive therapy: Assessment with a new self-rating instrument. *Biol Psychiatry* 1979;14:791–801. [PubMed: 497304]
52. Saykin, AJ. Dartmouth Medical School (Available from Author); 1992. Neurobehavioral Function and Activities of Daily Living Rating Scale (NBFADL-63 item version).
53. Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol* 1982;37:323–329. [PubMed: 7069156]
54. Santulli, RB.; Saykin, AJ.; Rabin, LA.; Wishart, HA.; Flashman, LA.; Pare, N., et al. Differential sensitivity of cognitive complaints associated with amnesic MCI: Analysis of patient and informant reports. Presented at the Alzheimer's Association International Conference on Prevention of Dementia; Washington, DC. Jun. 2005
55. Jorm AF. A short form of the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE): Development and cross-validation. *Psychol Med* 1997;24:145–153. [PubMed: 8208879]
56. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression scale: A preliminary report. *J Psychiatr Res* 1982;17:37–49. [PubMed: 7183759]
57. Saykin AJ, Wishart HA, Rabin LA, Santulli RB, Flashman LA, West JD, et al. Older adults with cognitive complaints show brain atrophy similar to that of amnesic MCI. *Neurology* 2006;67:834–842. [PubMed: 16966547]
58. Francis, WN.; Kucera, H. Frequency analysis of English usage: Lexicon and grammar. Boston: Houghton Mifflin Company; 1982.
59. Beglinger LJ, Gaydos B, Tangphao-Daniels O, Duff K, Kareken DA, Crawford J, et al. Practice effects and the use of alternate forms in serial neuropsychological testing. *Arch Clin Neuropsychol* 2005;20:517–529. [PubMed: 15896564]
60. Bird CM, Papadopoulou K, Ricciardelli P, Rossor MN, Cipolotti L. Test-retest reliability, practice effects and reliable change indices for the recognition memory test. *Br J Clin Psychol* 2003;42:407–425. [PubMed: 14633416]
61. Tabachnick, BG.; Fidell, LS. Using multivariate statistics. Vol. 3rd. New York: Harper Collins; 1996.
62. Rasmussen DX, Bylsma FW, Brandt J. Stability of performance on the Hopkins Verbal Learning Test. *Arch Clin Neuropsychol* 1995;10:21–26. [PubMed: 14588448]
63. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944. [PubMed: 6610841]
64. Portet F, Ousset PJ, Visser PJ, Frisoni GB, Nobili F, Scheltens Ph, et al. Mild cognitive impairment (MCI) in medical practice: A critical review of the concept and new diagnostic procedure. Report of the MCI Working Group of the European Consortium on Alzheimer's Disease. *J Neurol Neurosurg Psychiatry* 2006;77:714–718. [PubMed: 16549412]
65. Jessen F, Feyen L, Freymann K, Tepest R, Maier W, Heun R, et al. Volume reduction of the entorhinal cortex in subjective memory impairment. *Neurobiol Aging*. in press
66. Lactiz LH, Cullum CM. The Hopkins Verbal Learning Test and CVLT: A preliminary comparison. *Arch Clin Neuropsychol* 1998;13:623–628. [PubMed: 14590623]
67. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med* 2004;256:183–194. [PubMed: 15324362]
68. Ball CJ, Scott N, McLaren PM, Watson JP. Preliminary evaluation of a low cost video-conferencing (LCVC) system for remote cognitive testing of adult psychiatric patients. *Br J Clin Psychol* 1993;32:303–307. [PubMed: 8251960]
69. Mundt JC, Ferber KL, Rizzo M, Greist JH. Computer-automated dementia screening using a touch-tone telephone. *Arch of Intern Med* 2001;161:2481–2487. [PubMed: 11700161]
70. Buckwalter JG, Crooks VC, Petitti DB. A preliminary psychometric analysis of a computer-assisted administration of the Telephone Interview of Cognitive Status-Modified. *J Clin Exp Neuropsychol* 2002;24:168–175. [PubMed: 11992199]

71. Mintzer J, Nietert P, Costa K, Rust P, Hoernig K. Identifying persons with dementia by use of a caregiver telephone interview. *Am J Geriatr Psychiatry* 1998;6:176–179. [PubMed: 9581213]
72. Mueller, SG.; Weiner, MW.; Thal, LJ.; Petersen, RC.; Jack, CR.; Jagust, W., et al. Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). Vol. 1. *Alzheimer's & Dementia*; 2005. p. 55-56.

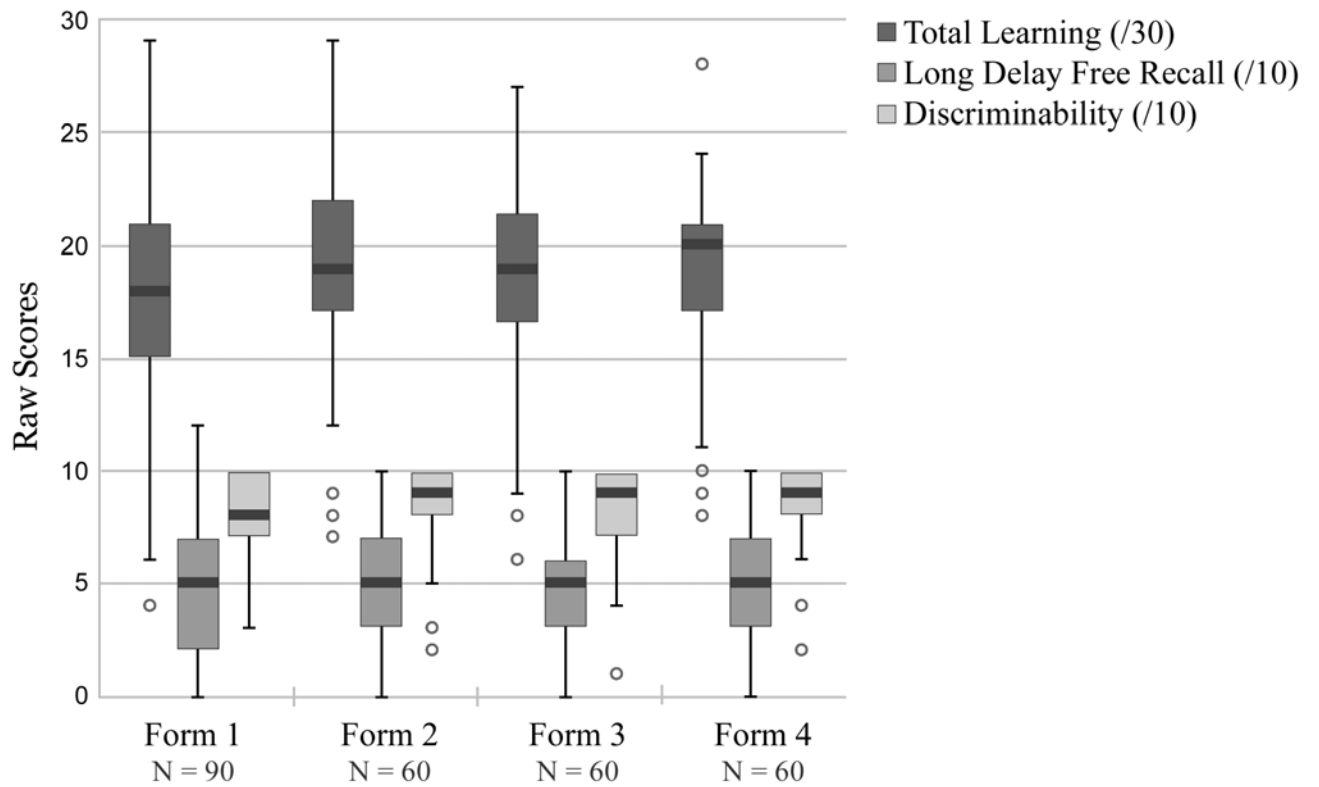


Fig. 1. MATS objective memory test scores by alternate test form. Box plots display medians, quartiles, and outliers of the four versions of the list-learning test, which did not differ significantly with respect to the learning, delayed recall, or recognition conditions.

Table 1

Form equivalencies on key variables

Variable	Analysis of Variance								Intraclass Correlation		
	Form 1		Form 2		Form 3		Form 4		p	r	p
	M	SD	M	SD	M	SD	M	SD			
Frequency	87.9	105.9	90.1	109.9	85.7	100.0	83.9	79.57	.99	.99	< .001
Letters	5.0	1.2	4.9	1.7	5.0	1.3	5.1	1.5	.83	.99	< .001
Syllables	1.5	.53	1.5	.53	1.5	.53	1.5	.53	1.0	1.0	< .001
Phonemes	4.0	1.2	3.8	1.3	4.3	1.4	3.9	.99	.82	.82	< .001

Table 2
Participant demographics and neuropsychological test data for validation study

Characteristic	Participant Group*						p	Group Differences
	HC N = 43		CC N = 38		MC IN = 39			
	M	SD	M	SD	M	SD		
Age, yrs.	70.56	5.18	73.45	5.91	72.10	6.34	.090	---
Education, yrs.	16.77	2.50	16.45	2.88	16.28	3.03	.727	---
Gender (M, F)	12, 31	---	16, 22	---	17, 22	---	.266	---
ANART [†]	122.56	4.54	121.32	6.46	120.05	7.33	.189	---
GDS (/26) [‡]	1.49	2.22	4.18	3.95	3.69	3.11	<.001	HC < CC, MCI
MMSE, Total Score (/30) [§]	29.12	1.03	28.92	1.19	26.64	2.08	<.001	HC, CC > MCI
DRS-2								
Total Score (/144)	141.02	1.96	140.95	2.70	137.00	4.26	<.001	HC, CC > MCI
Memory Subscale (/25)	23.93	1.06	24.13	1.19	21.59	2.92	<.001	HC, CC > MCI
CVLT-II[#]								
Learning (/80)	49.23	7.28	45.79	8.88	31.95	6.48	<.001	HC, CC > MCI
Short Delay Free Recall (/16)	11.37	2.11	10.13	2.61	5.33	2.53	<.001	HC, CC > MCI
Long Delay Free Recall (/16)	12.05	2.22	10.61	2.78	5.69	2.55	<.001	HC > CC > MCI
Discriminability (/4)	3.21	0.61	2.94	.724	1.82	.809	<.001	HC, CC > MCI
WMS-III^{**}								
Logical Memory								
Immediate Recall (/75)	48.16	7.37	45.21	8.52	33.87	9.50	<.001	HC, CC > MCI
Delay Recall (/50)	31.33	6.25	27.58	6.82	18.28	8.37	<.001	HC, CC > MCI
Visual Reproduction								
Immediate Recall (/104)	78.19	11.77	74.95	12.62	63.18	17.76	<.001	HC, CC > MCI
Delay Recall (/104)	54.07	20.4	41.50	16.25	28.79	17.37	<.001	HC > CC > MCI

* HC = healthy control; CC = cognitive complaint; MCI = mild cognitive impairment

[†]ANART = American National Adult Reading Test, estimate of baseline Verbal IQ

GDS = Adjusted Geriatric Depression Scale (n / 26 non-cognitive items)

§ MMSE = Mini Mental State Exam

¶ DRS-2 = Dementia Rating Scale-2

CVLT-II = California Verbal Learning Test, Second Edition

** WMS-III = Wechsler Memory Scale, Third Edition

Table 3

Mean MATS* scores by participant group

MATS Variable	Participant Group †						p, Effect Size‡	Group Differences
	HC N = 43		CC N = 38		MCI N = 39			
	M	SD	M	SD	M	SD		
Subjective Memory Test								
Total Score (/ 10)	2.30	1.63	4.63	1.81	5.28	2.19	< .001, .33	CC, MCI > HC
Objective Memory Test								
Learning (/ 30)	20.74	3.59	17.95	3.03	16.15	3.21	< .001, .26	HC > CC > MCI
Long Delay Free Recall (/ 10)	6.84	1.77	5.32	2.00	3.38	1.83	< .001, .38	HC > CC > MCI
Discriminability (/ 10)	9.18	1.03	8.56	1.37	7.20	1.62	< .001, .28	HC > CC > MCI§

* MATS = Memory and Aging Telephone Screen

† HC = healthy control; CC = cognitive complaint; MCI = amnesic mild cognitive impairment

‡ Partial Eta squared

§ Denotes results of Mann-U Whitney tests

Table 4
Participant demographics and MATS* objective memory test performance for alternate test form study

Characteristic	Participant Group [†]						p	Group Differences
	HC N = 25	CC N = 25	MCI N = 25	AD N = 15	M	SD		
	MATS Objective Memory Test Performance (Form 1)							
Learning (/ 30)	20.92	19.60	16.28	12.47	5.06	<.001	HC, CC, > MCI > AD	
Long Delay Free Recall (/ 10)	7.16	5.64	3.40	2.20	3.05	<.001	HC, CC, > MCI, AD	
Discriminability (/ 10)	9.19	9.17	6.91	5.25	1.88	<.001	HC, CC, > MCI > AD [‡]	

* MATS = Memory and Aging Telephone Screen

[†] HC = healthy control; CC = cognitive complaint; MCI = amnesic mild cognitive impairment; AD = probable mild Alzheimer's disease

[‡] Denotes results of Mann-U Whitney tests

Table 5
Mean MATS objective memory test scores by test form and condition (repeated measures design)

Test Forms	MATS* Objective Memory Test Variable											
	Total Learning (/ 30)			Long Delay (/ 10)			Discriminability (/ 10)			p		
	Mean	SD	p	Mean	SD	p	Mean	SD	p	Mean	SD	p
N = 30												
1	17.63	4.28	.18	4.43	2.64	.232	7.43	2.18	.06			
2	19.30	4.69		5.23	2.88		8.40	2.32				
3	18.67	5.01		4.40	2.61		7.73	2.40				
N = 30												
1	18.00	5.06	.57	5.33	3.19	.840	8.27	1.84	.41			
3	18.40	4.55		4.97	2.53		8.50	1.70				
4	18.80	4.95		5.03	2.74		8.73	1.57				
N = 30												
1	17.93	4.93	.35	4.83	2.38	.922	8.17	2.18	.23			
2	19.17	4.36		4.67	2.54		8.47	2.11				
4	19.17	3.00		4.83	2.18		8.07	2.08				

* MATS = Memory and Aging Telephone Screen

† p values refer to repeated measures ANOVAs for each test condition (i.e., learning, delayed recall, and recognition discriminability) for the 30 participants who completed each combination of test versions: Forms 1, 2, 3; Forms 1, 3, 4; or Forms 1, 2, 4.

Box 1

Criteria used to classify study participants

	HC	CC	MCI	AD
(1) abnormal memory performance*			✓	✓
(2) significant memory complaints, corroborated by an informant [†]		✓	✓	✓
(3) relatively preserved general cognitive functioning	✓	✓	✓	
(4) generally normal activities of daily living	✓	✓	✓	
(5) no dementia	✓	✓	✓	
(6) no depression or other major psychiatric disorder	✓	✓	✓	✓

* At least 1.5 SDs below the mean established for age- and education-matched controls on at least one standardized test of episodic memory (please see Saykin et al., 2006 for more detail).

[†] Endorsed at least 20% of possible cognitive complaints across all inventories or complaints deemed significant by clinical consensus.

Box 2
Overview of MATS* informational sections

Basic demographic information

name, telephone number, address, name of physician/date of last visit
date of birth, handedness, yrs of education, parents' yrs of education
profession/occupation, primary language, height, weight
referral source, factors leading to participation (e.g., concern about memory, helping research)
name and relation of collateral informant

Current and past medical and psychiatric history

overview of health concerns, surgeries, neuroimaging, hearing, vision, medications/dosages
vitamins, nutritional supplements, hypertension, current blood pressure, cholesterol
cancer (and consequent treatments), neurological diseases, headaches/migraines, respiration
arthritis, diabetes, sleep, appetite, balance, incontinence
alcohol, drug, or nicotine use, depression, anxiety, psychosis, learning or attention difficulties
family medical and psychiatric history (particularly with regard to memory loss)

Pre-MRI metal screen

metal in body or eyes, pacemaker, aneurysm clip, neurostimulator
dental braces, non-removable retainer, bone fractures treated with screws, rods, pins
cochlear, otologic, or ear implant
tattoos or permanent eye liner, claustrophobia

Description of study

overview of study goals and assessment techniques (e.g., neuropsychological evaluation, hearing
test, blood test, hearing test, functional and structural MRI)
time requirements, stipend

* MATS = Memory and Aging Telephone Screen

Box 3

Overview of MATS* subjective memory test

Item	Response †		
1. Remembering conversations	N	Y	DK
2. Word finding (e.g., coming up with the right word, tip-of-tongue phenomenon)	N	Y	DK
3. More difficulty remembering names (relative to age peers)	N	Y	DK
4. More difficulty learning new information (relative age to peers)	N	Y	DK
5. Finding everyday objects	N	Y	DK
6. Remembering what you entered a room to do	N	Y	DK
7. Memory now compared to the best it's ever been	N	Y	DK
	same	worse	
8. Memory compared to age peers	N	Y	DK
9. Effect on daily activities (e.g., social, vocational, recreational)	N	Y	DK
10. Progression over time (Duration)		Years _____ ‡	
		Months _____	
11. Evaluation or treatment for memory loss (specify evaluation type and any medications)	N	Y	DK
12. Other problems with memory (in addition to those already mentioned)	N	Y	DK
<i>Additional Question: Age of onset of memory problems and duration/course</i>	N	Y	DK

* MATS = Memory and Aging Telephone Screen; complete test protocols available upon request.

† Responses receive a score of 0 for "no" or "same" and 1 point for "yes" or "worse." Participants are encouraged to provide valid, scorable responses for all items. Total score is derived by summing the points attained (range = 0 - 12).

‡ This item is not scored.

Box 4

MATS* objective memory test: alternate forms of list learning stimuli†

Form 1	Form 2	Form 3	Form 4
hat	sky	rain	beach
engine	baseball	wagon	flower
dancer	farmer	movie	cousin
fort	snow	salt	star
chair	rock	bird	gas
garden	kitchen	forest	mountain
apron	apple	orange	wallet
crystal	lion	costume	sugar
face	car	foot	book
pill	brick	soap	pipe

* MATS = Memory and Aging Telephone Screen

† Test protocols, including recognition stimuli, are available upon request.