OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

The Indiana Faces in Places Test: Preliminary Findings on a New Visuospatial Memory Test in Patients with Mild Cognitive Impairment

Leigh J. Beglinger[a],*, Kevin Duff[a], David J. Moser[a], Stephen A. Cross[a], David A. Kareken[b]

[a]Department of Psychiatry, University of Iowa College of Medicine, Iowa City, IA, USA
[b]Department of Neurology, Indiana University School of Medicine, IN, USA

Abstract

Memory assessment is an important component of a neuropsychological evaluation, but far fewer visual than verbal memory instruments are available. We examined the preliminary psychometric properties and clinical utility of a novel, motor-free paper and pencil visuospatial memory test, the Indiana faces in places test (IFIPT). The IFIPT and general neuropsychological performance were assessed in 36 adults with amnestic mild cognitive impairment (aMCI) and 113 older adults with no cognitive impairment at baseline, 1 week, and 1 year. The IFIPT is a visual memory test with 10 faces paired with spatial locations (three learning trials and non-cued delayed recall). Results showed that MCI participants scored lower than controls on several variables, most notably total learning ($p < .001$ at all three time points), delayed recall (baseline $p = .03$, 1 week $p < .001$, 1 year $p < .001$), and false-positive errors (range $p = .03$ to <0.001). The IFIPT showed similar test–retest reliability at 1-week and 1-year follow-up to other neuropsychological tests ($r = 0.71–0.84$ for MCI and $0.53–0.72$ for controls). Diagnostic accuracy was modest for this sample (areas under the receiver operating characteristic curve between 0.64 and 0.66). Preliminary psychometric analyses support further study of the IFIPT. The measure showed evidence of clinical utility by demonstrating group differences between this sample of healthy adults and those with MCI.

Keywords: Mild cognitive impairment; Visual memory; Face memory; Test–retest reliability

Introduction

Memory assessment is a core component of a neuropsychological examination. Deficits in this domain are the central symptom in a variety of neurological and psychiatric conditions, such as dementia, amnestic disorders, alcohol abuse, stroke, encephalitis, traumatic brain injury, depression, and epilepsy (Brown, Tapert, Granholm, & Delis, 2000; Butters, Delis, & Lucas, 1995; Butters, Wolfe, Martone, Granholm, & Cermak, 1985; Lezak, 1979; Paulsen et al., 1995). Further, the specific type and pattern of memory impairment provides important clinical information that aids in diagnosis, lateralization of lesions, and predictions about functional improvement or decline. Visual memory impairment is an important symptom for assessment, but the library of clinically useful visual memory measures has lagged behind its verbal counterpart (Barr, 1997; Mapstone, Steffenella, & Duffy, 2003).

There are several well-validated and commonly used visual memory tests, including design reproduction and recall tests (e.g., Wechsler memory scale visual reproduction, Brief visuospatial memory test-revised, Benton visual retention test, and Rey–Osterrieth complex figure test), visual recognition tests (e.g., continuous visual memory test), and visual learning tests (e.g., 7/24 spatial recall test; Lezak, Howieson, Loring, Hannay, & Fischer, 2004). Each of these visual memory tests has strengths and weaknesses. Ideally, a visual memory test should be relatively easy to administer, inexpensive, and portable and has a good range of difficulty so that it can be used in patients with both subtle and distinct impairments. Additionally,

---

* Corresponding author at: Department of Psychiatry, University of Iowa College of Medicine, Iowa City, IA 52242-1000, USA.
E-mail address: leigh-beglinger@uiowa.edu (L.J. Beglinger).

it is advantageous if the test has multiple trials to examine learning, which is relevant in many patient populations (e.g., degenerative diseases, traumatic brain injury). Many of the common memory tests meet these requirements. It is also beneficial for many neurological populations if the visual test is motor-free so that fine and gross motor impairments do not confound the assessment of memory. Unfortunately, most of the traditionally used visual memory measures require a motor response, which limits the types of patients with which they can be used. Finally, over the last two decades, there has been a burgeoning interest in evaluating the ecological validity of neuropsychological tests. Tests that provide a closer approximation to everyday tasks are appealing (Larrabee & Crook, 1988). Many visual memory measures utilize abstract stimuli in an attempt to obviate verbalization strategies during encoding, but those types of stimuli are somewhat removed from real-world tasks. In the present study, we present preliminary results for a novel visual memory test that addresses many of the issues raised above: It is motor-free, it uses faces as stimuli which may provide greater ecological validity than abstract figures, it contains learning trials and an incidental recall trial and it is portable.

Visuospatial deficits, including visual learning and memory, have been found on a continuum from normal aging to dementia (Mapstone et al., 2003). These deficits appear early in the dementing process and have been shown to discriminate subtle cognitive impairment from Alzheimer's disease (AD; Alescio-Lautier et al., 2007; Blackwell et al., 2004; Fowler, Saling, Conway, Semple, & Louis, 2002), making them a significant target for further research. Additionally, there is evidence that visuospatial memory measures may be more sensitive to age-related cognitive changes than verbal measures, which is an added benefit when subtle differences are to be detected (Jenkins, Myerson, Joerding, & Hale, 2000). Although the concept of amnestic mild cognitive impairment (aMCI) as a formal and well-defined diagnostic entity remains controversial (Davis & Rockwood, 2004; Winblad et al., 2004), there is a recognized clinical and research need for early identification of patients with memory impairment. Recent research suggests that visuospatial memory performance may be a sensitive predictor of decline (Fowler, Saling, Conway, Semple, & Louis, 1995, 1997, 2002; Griffith et al., 2006). Face learning and recognition are strongly localized to right temporal lobe regions and have been shown to be highly sensitive to early impairment in the elderly due to the complexity of facial stimuli (Dade & Jones-Gotman, 2001; Haxby, Hoffman, & Gobbini, 2000; Werheid & Clare, 2007). Thus, it is not surprising that these types of tests would elicit deficits in early AD and aMCI. However, the particular visuospatial memory tests used in the majority of this research have limitations. The paired associate learning test from the Cambridge automated neuropsychological test assessment battery (CANTAB; Morris, Evenden, Sahakian, & Robbins, 1987) was used in four studies. Other non-paper and pencil tests requiring presentation of visual stimuli with a projector have also been used (Mapstone et al., 2003). Although computerized assessment batteries are advantageous in a number of ways (e.g., ease of administration and more detailed measurement of reaction time), they are not without limitations. Most automated batteries are not comprehensive and require supplementation with additional tests. Patient rapport, effort, and compliance can be more problematic, and cost of the software and hardware is often prohibitive. Additionally, some of the tests have ceiling effects (e.g., the dementia rating scale and the CANTAB paired associate learning) which may reduce their utility in discriminating early impairments from normal performance, a critical distinction in MCI (Fowler et al., 2002; Mapstone et al., 2003).

The purpose of the current study was to conduct a preliminary evaluation of the psychometric properties of a new memory measure. As part of this evaluation, we examined the clinical utility of the Indiana faces in places test (IFIPT; developed by Beglinger and Kareken). Specifically, test–retest reliability, sensitivity, specificity, positive and negative predictive powers, and area under the receiver operating characteristic (ROC) curve were calculated in a sample of patients with MCI and a healthy elderly comparison group.

Materials and Methods

Participants

One hundred and forty-nine community-dwelling older adults (aged 65 years and older) served as participants for this study. The sample was recruited at a variety of living facilities for senior citizens (e.g., retirement communities, independent living facilities) and senior centers through local talks and advertisements in the community and surrounding areas. Exclusion criteria included significant history of major neurological (e.g., traumatic brain injury, stroke, dementia) or psychiatric illness (e.g., schizophrenia, bipolar disorder), possible mental retardation based on a wide-range achievement test-3 (WRAT-3; Wilkinson, 1993) reading <70 or current depression (either self-report or 30-item geriatric depression scale [GDS; Yesavaga et al., 1983] of >15). All data were reviewed by two neuropsychologists (KD and LJB), and participants were classified into two groups, either aMCI or normal comparison (NC), using existing criteria (Petersen et al., 1999) and by expert consensus review as follows. To be classified as aMCI, all participants had to complain of memory problems (i.e., self-reported as yes/no during an interview). These aMCI participants had to have objective memory deficits on the Hopkins verbal learning

test-revised (HVLT-R; Brandt & Benedict, 2001) with delayed recall falling 1.5 *SD* or more below normative average (i.e., standard score <78). Second, overall cognition had to be otherwise intact (i.e., age-corrected repeatable battery for the assessment of neuropsychological status [RBANS; Randolph, 1998] Total Scale score greater than 1.5 standard deviations below average). This cut-off point of 1.5 *SD* below average is common in MCI research. NC participants were not excluded if they had memory complaints, as cognitive complaints are fairly common in older adults and do not necessarily correspond to objective deficits. These NC participants did not present with objective memory deficits (i.e., HVLT-R delay recall scores of better than 1.5 *SD* below average). Similar to the aMCI participants, individuals classified as NC had intact overall cognition. Of the 149 individuals, 36 were classified with aMCI and 113 as NC. No one was classified as demented (i.e., impairments in memory and other cognitive domains and activities of daily living). Functional status was assessed during a telephone screening interview in which both participants and a collateral were asked about ADL's and IADL's. Demographic and baseline assessment scores are presented in Table 1. All participants provided written informed consent for the study and were financially compensated for their time. Additional details about the test performances of the two groups are presented elsewhere. This study is part of a larger, ongoing study evaluating practice effects in MCI (Duff et al., 2008).

Procedures

The research protocol and all study procedures were approved by the University of Iowa Institutional Review Board. Participants were screened for enrollment using a brief clinical interview, RBANS form A, WRAT-3 reading subtest, and the GDS. The clinical interview assessed relevant demographic information, medical and psychiatric history, presence of memory complaints, and report of activities of daily living. A collateral source (e.g., spouse, adult child, close friend) completed a similar interview to corroborate the reports by the participant. After meeting eligibility criteria for enrollment, participants were assessed at a baseline visit that consisted of a 60-min neuropsychological screening battery to obtain measures of functioning in memory, attention, psychomotor speed, and executive functions. Eligible participants were then re-assessed a week later using the same battery to explore practice effects and test–retest reliability. All assessments were conducted by a trained research assistant or by one of the neuropsychologists (LJB or KD). Alternate forms of the testing battery were not used so that test–retest analyses across standard measures would be comparable with the IFIPT.

A subset of the sample has also been re-evaluated with the same battery after 1 year. The 1-year follow-up evaluations are ongoing and the following samples were available for this analysis: 21 individuals classified as MCI at baseline and 92 individuals classified as NC at baseline.

Measures

Measures in the neuropsychological battery include: Symbol digit modalities test (SDMT; Smith, 1991); HVLT-R; brief visuospatial memory test-revised (BVMT-R) (Benedict, 1997); controlled oral word association test (COWAT) and animal fluency (Benton, Hamsher, Varney, & Spreen, 1983); modified mini-mental state examination (3MS; Teng & Chui, 1987); temporal and spatial orientation items; IFIPT; and trail making test parts A and B (Reitan, 1958). All these measures were administered and scored as described in their respective test manuals.

Table 1. Demographic characteristics and baseline memory scores in participants with amnestic MCI and intact cognition ($N = 149$)

Demographics	aMCI ($N = 36$) [mean (*SD*)]	NC ($N = 113$) [mean (*SD*)]
Age (years)	79.47 (8.15)	77.96 (7.64)
Education (years)	14.94 (2.70)	15.66 (2.61)
Sex		
Men	11 (31%)	18 (16%)
Women	25 (69%)	95 (84%)
WRAT-3 reading standard score	107.42 (5.94)	108.11 (6.17)
HVLT-R delay	66.26 (7.89)***	100.73 (12.20)
BVMT-R delay	71.72 (16.15)***	92.11 (19.12)
RBANS delayed memory index	93.22 (14.03)***	103.16 (11.49)

Notes: Standard scores are presented for each cognitive measure. MCI = mild cognitive impairment; WRAT-3 = wide range achievement test-3; RBANS = repeatable battery for the assessment of neuropsychological status; BVMT-R = brief visuospatial memory test-revised; HVLT-R = Hopkins verbal learning test-revised.
***$p < .001$.

IFIPT Administration and Scoring

The IFIPT is a novel paper-and-pencil visuospatial memory test that examines visual learning and memory. The test is comprised of 10 target black-and-white faces (as developed by the University of Pennsylvania Brain Behavior Laboratory) paired with 10 spatial locations represented by boxes on a page. The visual array is shown in Fig. 1. Faces were chosen to be representative of the gender, age, and ethnic diversity of the USA (as a whole). The faces were sized so that little other identifying information (e.g., clothing) was visible and the actors posed with neutral facial expressions. Through pilot testing, 10 faces were chosen as targets because this number provided enough difficulty to avoid both floor and ceiling effects during the learning trials. For each target face, three matching foils were chosen for the recognition trials described below.

The test consists of three learning trials, two immediate recognition trials, and one non-cued delay recognition trial. In the first learning trial (i.e., Trial 1), participants are shown the 10 target faces, one at a time on the location grid, at the rate of one face-grid pairing every 2 s. Immediately following this learning trial, the first recognition trial is administered. Participants view single faces on a blank sheet of paper and are asked whether they had previously studied the face on the preceding learning trial. There are 20 faces in the recognition trial (10 targets and 10 foils, matched in gender, age, and race). If the participant indicated that he/she has seen the face before during the learning trial, then the administrator presents the participant with the
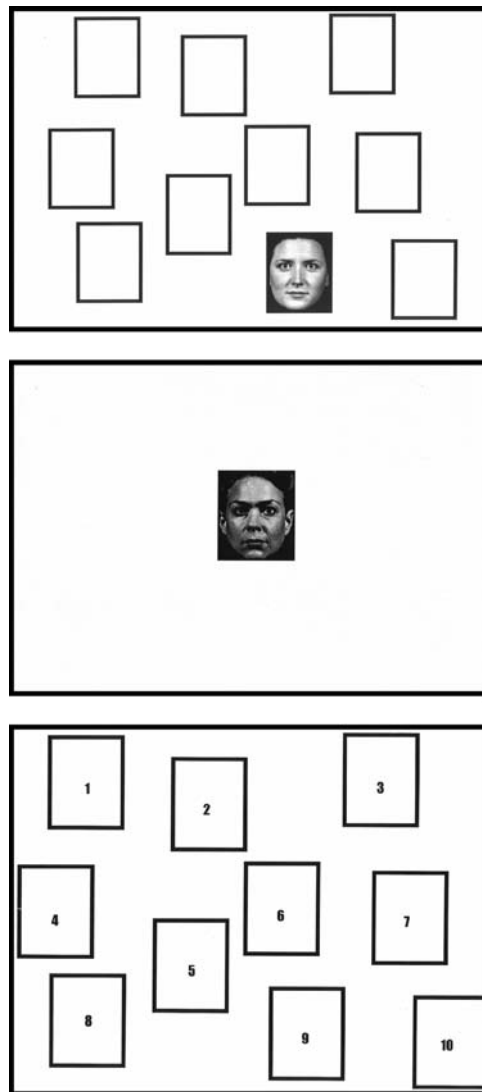

Fig. 1. (A) An example of one Indiana faces in places test target face stimulus from a learning trial; (B) one foil stimulus from a recognition trial; (C) location grid for recognition trials.

location grid and he/she is asked to identify in which spatial location the face was presented previously. Following completion of the first recognition trial, a second learning trial (i.e., Trial 2) is administered. There is no recognition trial for Trial 2 as it added unnecessary administration time and was not a useful additional score in pilot testing. A third learning trial (i.e., Trial 3) immediately follows. In the Trial 3 recognition trial, the 10 targets and 10 novel foils are presented one at a time on individual sheets of paper and the participant is asked if he/she has seen this face on the learning trials. Like in the Trial 1 recognition trial, if a face is identified as previously viewed, then the location grid is provided and the participant is asked to identify the spatial location of that face. A delayed recognition trial is administered approximately 30 min after the third recognition trial. Participants are again given the location grid and are asked to identify the target faces and associated spatial locations among a set of the 10 targets and 10 new foils. Total administration time, excluding the interval between learning and delayed recall, is approximately 15 min.

Each recognition trial (i.e., Trial 1, Trial 3, delay) yields three scores: Number correct, location hits, and location false positives. For each face in the recognition trial, participants are first asked "Have you seen this face before—yes or no?" Number correct is number of correct responses to this question, which has a maximum of 20 (10 targets and 10 foils). If a face is positively identified (e.g., "yes, I've seen that face on the learning trial"), then the participant must identify the correct spatial location on the grid. Location hits are the number of correctly placed faces on the grid, with a maximum of 10 (10 target locations). A total learning score was calculated by summing location hits from Trials 1 and 3. Location false positives are the total number of faces that are incorrectly placed on the grid, with a maximum of 20.

Statistical Analyses

To examine test–retest reliability, Pearson's correlations were calculated between the baseline, 1-week, and 1-year assessments on the IFIPT total learning and delay hits. Paired-sample *t*-tests were used to examine the change on the IFIPT primary variables between baseline and 1 week and baseline and the 1-year assessment for each group, as well as on the HVLT-R and BVMT-R for comparison. To examine convergent validity, the IFIPT total learning and delay hits at baseline were compared with performances on other memory measures (RBANS figure recall, BVMT-R total recall and delay recall, HVLT-R total recall and delay recall) with Pearson's correlations. Divergent validity was examined by calculating correlations between IFIPT variables and non-memory measures from the battery (trail making test, COWAT, animal fluency, WRAT-3 reading, GDS). To examine validity to detect group differences, *t*-tests were used to compare the NC and MCI groups on the 10 IFIPT variables. The α-value was set at 0.05. Sensitivity, specificity, and positive and negative predictive powers were calculated at baseline using the base rates of MCI within the current sample; area under the curve using ROC curves was used to determine diagnostic classification accuracy.

Results

Patient Characteristics

The NC and aMCI participants were comparable on all demographic variables, including age ($p = .31$), years of education ($p = .16$), estimated premorbid verbal skills (i.e., WRAT-3 reading ($p = .56$), and gender ($p = .54$). The current sample was reflective of the rural Midwest and unintentionally not ethnically diverse (100% Caucasian).

Test–Retest Reliability

IFIPT total learning at baseline significantly correlated with total learning at the 1-week follow-up visit ($r[144] = 0.74$, $p < .001$) and 1-year follow-up visit ($r[106] = 0.60$, $p < .001$) for all participants. Delay hits on the IFIPT at baseline also significantly correlated with delay hits at 1 week ($r[144] = 0.72$, $p < .001$) and 1 year ($r[108] = 0.59$, $p < .001$). When the groups were separated into MCI and controls, correlations were higher in the MCI group. In the MCI participants, IFIPT total learning at baseline significantly correlated with total learning at the 1-week follow-up visit ($r[34] = 0.83$, $p < .001$) and 1-year follow-up visit ($r[21] = 0.84$, $p < .001$). Delay hits on the IFIPT at baseline also significantly correlated with delay hits at 1 week ($r[34] = 0.70$, $p < .001$) and 1 year ($r[20] = 0.73$, $p < .001$). In controls, IFIPT total learning at baseline significantly correlated with total learning at the 1-week follow-up visit ($r[107] = 0.68$, $p < .001$) and 1-year follow-up visit ($r[87] = 0.54$, $p < .001$). Delay hits on the IFIPT at baseline also significantly correlated with delay hits at 1 week ($r[110] = 0.72$, $p < .001$) and 1 year (r$[88] = 0.53$, $p < .001$).

Change from Baseline to 1 Week

In the MCI group of individuals who completed both the baseline and the 1-week follow-up, results of paired-sample *t*-tests indicated that performance improved significantly on the learning score for all three memory tests after 1 week (IFIPT total learning, BVMT-R total recall, and HVLT-R total recall, all $p < .001$), as well as the delay score for all three tests (IFIPT delay hits, BVMT-R delayed recall, and HVLT-R delayed recall, all $p < .001$), presumably reflecting practice effects. *T*-tests in control participants also revealed that controls performed significantly better on all three memory learning and delay tasks between baseline and 1 week (IFIPT, BVMT-R, and HVLT-R, all $p < .001$).

Change from Baseline to 1 Year

On the learning scores, MCI participants performed significantly better after 1-year compared with baseline on the IFIPT total learning score ($t = -2.71$, $p = .01$), but not on the BVMT-R total recall ($t = 1.55$, $p = .13$) or HVLT-R total recall ($t = -1.51$, $p = .15$). On the delayed recall scores, there was no improvement over 1 year on the IFIPT delay hits ($t = -0.63$ $p = .54$) or BVMT-R delay ($t = -0.74$, $p = .47$), but there was significant improvement on the HVLT-R delay ($t = -2.97$, $p = .007$). Among the controls, significant improvements were observed across 1 year on the IFIPT total learning ($t = -4.84$, $p = .001$) and BVMT-R total recall ($t = -2.82$, $p = .006$), but not on the HVLT-R total recall ($p = .17$). The delay scores for IFIPT ($t = -4.31$, $p = .001$) and BVMT-R ($t = -2.83$, $p = .006$) again showed significant improvements in controls across 1 year, but not the HVLT-R ($p = .60$).

Validity

Pearson's correlations between the two main IFIPT variables and other neuropsychological tests are reported in Table 2 for the two participant groups separately. Correlations in the control group revealed that baseline IFIPT total learning and delay hits correlated most strongly with BVMT-R total recall ($r = 0.42$ and 0.50, $p < .001$) and BVMT-R delay recall ($r = 0.43$ and 0.49, $p < .001$). Both IFIPT scores were also correlated at $p < 0.01$ with RBANS figure recall, HVLT-R total and delayed recall, trail making test parts A and B, and animal fluency. IFIPT total learning and delay hits at baseline correlated with each other at $r = 0.74$. In the MCI group, again baseline IFIPT total learning and delay hits correlated most strongly with BVMT-R total recall ($r = 0.71$ and 0.64, $p < .001$) and BVMT-R delay recall ($r = 0.57$ and 0.46, $p < .001$). IFIPT scores

Table 2. Pearson's correlations between the IFIPT and neuropsychological variables at baseline for amnestic MCI group and intact cognition group

	aMCI ($N = 36$)		NC ($N = 111$)	
	IFIPT total learning	IFIPT delay hits	IFIPT total learning	IFIPT delay hits
IFIPT total learning	—	.737***	—	.735***
IFIPT delay hits	.737***	—	.735***	—
RBANS picture naming	−.001	.108	.170	.188*
RBANS figure recall	.536**	.237	.347***	.285**
BVMT-R total recall	.706***	.636***	.422***	.495***
BVMT-R delayed recall	.570***	.462**	.434***	.489***
HVLT-R total recall (Trials 1–3)	.345*	.271	.302**	.325**
HVLT-R delayed recall	.114	.−.140	.319**	.445***
COWAT	.068	.099	.008	.082
Animal fluency	.450**	.203	.357***	.374***
TMTA seconds	−.474**	−.448**	−.327**	−.308**
TMTB seconds	−.541**	−.450**	−.263**	−.345***
SDMT correct	.555***	.525**	.178	.238*
WRAT-3 reading	−.046	−.214	.135	.167
RBANS digit span	.250	.363*	.090	.118
GDS	−.134	−.171	.113	.117

Notes: IFIPT = Indiana faces in places test; MCI = mild cognitive impairment; RBANS = repeatable battery for the assessment of neuropsychological status; BVMT-R = brief visuospatial memory test-revised; HVLT-R = Hopkins verbal learning test-revised; WRAT-3 = wide range achievement test-3; COWAT = controlled oral word association test; TMT = trail making test (reverse coded); SDMT = symbol digit modalities test; GDS = geriatric depression scale.
*Correlation is significant at the $p < .05$ level (two-tailed).
**Correlation is significant at the $p < .01$ level.
***Correlation is significant at the $p < .001$ level.

were also associated at $p < .01$ with Trails A and B and SDMT. The IFIPT total learning was also associated with other neuropsychological tests, such as RBANS figure recall, HVLT-R total recall, and animal fluency. Neither IFIPT total learning nor delay hits correlated with estimated premorbid verbal skill (WRAT-3 reading) or baseline GDS score in either the MCI group or controls.

Diagnostic Accuracy

Results of independent samples *t*-tests indicated that the MCI group performed significantly below the NC group on the IFIPT Trial 3 hits at baseline ($p = .001$), Trial 3 false positives ($p = .03$), total learning ($p = .002$), delay correct ($p = .02$), and delay hits ($p = .03$). The MCI group performed significantly below the controls on all 10 IFIPT variables except Trial 3 correct at 1 week and Trial 1 hits at 1 year, as shown in Table 3.

Sensitivity, specificity, and positive and negative predictive powers are all presented in Table 4. ROC curves at baseline are presented in Fig. 2. The area under the curve for IFIPT total learning was 0.67 and for IFIPT delay hits was 0.63.

Discussion

Our findings lend preliminary support to the IFIPT as a test of visuospatial learning and memory. We found the IFIPT to have acceptable test–retest reliability, adequate convergent validity with other memory measures, and to discriminate between patients with MCI and older adults with no cognitive impairment. An additional benefit is that the IFIPT is a nonverbal measure that did not correlate with estimated premorbid verbal skills or depressive symptoms, which can be a useful addition to a battery. It is also a motor-free visual memory test, which is an advantage in assessing patients with motor deficits. Finally, the IFIPT was of sufficient difficulty to avoid ceiling effects on the main variables at baseline; no one in the MCI group obtained a perfect score on Trial 1, Trial 3, or delay hits and only one person in the control group obtained a perfect score on only one of those variables (Trial 3 hits), which supports the clinically utility of the IFIPT in distinguishing subtle changes in visual memory. Tests that are capable of identifying patients with mild memory impairment are crucial for both clinical care and in research as treatment trials are being extended downward from dementia to pre-dementia conditions.

Across 1 week, the two primary outcome measures of the IFIPT demonstrated adequate stability coefficients of 0.68 and 0.72 in controls and slightly higher correlations in the MCI group (0.83 and 0.71). Over a 1-year interval, correlations were 0.47 and 0.53 in controls and 0.84–0.73 in MCI. Although the 1-year retest correlations in the control group are psychometrically modest, these correlations are comparable with other studies of neuropsychological measures with ranges in the upper

Table 3. IFIPT scores at baseline, 1-week, and 1-year follow-up in aMCI and NC participants

IFIPT variables	Baseline, mean (*SD*), range		One-week follow-up, mean (*SD*), range		One-year follow-up, mean (*SD*), range	
	aMCI (*n* = 36)	NC (*n* = 111)	aMCI (*n* = 34)	NC (*n* = 113)	aMCI (*n* = 21)	NC (*n* = 92)
Trial 1 correct	11.94 (2.03), 9–16	12.11 (2.09), 7–18	14.44* (2.61), 8–20	15.47(2.58), 9–20	12.57** (2.11), 9–16	14.54 (2.71), 8–20
Trial 1 hits	1.75 (1.52), 0–5	2.21 (1.63), 0–7	4.00* (2.56), 0–9	5.10 (2.45), 0–10	2.24 (1.64), 0–5	2.97 (1.85), 0–8
Trial 1 false positives	9.69 (3.64), 4–20	9.49 (3.49), 2–18	8.71* (3.59), 1–17	7.32 (3.43), 0–17	10.33* (4.26), 3–19	8.64 (2.87), 2–17
Trial 3 correct	16.06* (2.11), 11–20	16.54 (2.52), 9–20	16.71 (2.46), 10–20	17.57 (2.18), 10–20	16.10** (2.86), 9–20	17.82 (2.40), 10–20
Trial 3 hits	3.89** (2.38), 0–9	5.30 (2.07), 1–10	5.21*** (2.35), 1–9	6.88 (2.25), 1–10	3.95*** (2.27), 0–7	6.21 (2.25), 1–10
Trial 3 false positives	7.22* (3.34), 1–13	5.83 (3.23), 0–14	6.44** (3.65), 1–15	4.32 (3.65), 0–19	6.90** (4.17), 0–16	4.46 (3.62), 0–18
Total learning	5.64** (3.52), 1–14	7.59 (3.09), 1–16	9.21** (4.62), 2–18	12.01 (4.27), 2–20	6.19** (3.70), 0–12	9.21 (3.59), 2–18
Recognition delay correct	14.80* (2.43), 9–20	15.92 (2.52), 9–20	15.29** (2.53), 10–20	16.88 (2.35), 9–20	15.20** (2.63), 10–19	17.01 (2.23), 11–20
Recognition delay hits	3.40* (2.19), 1–9	4.33 (2.16), 0–9	4.74** (2.25), 1–9	6.26 (2.44), 0–10	3.55** (2.46), 0–8	5.41 (2.21), 0–9
Recognition delay false positives	7.80 (3.22), 2–14	6.84 (3.29), 0–17	7.74** (3.16), 1–13	5.46 (3.55), 0–19	8.15** (4.78), 1–20	5.56 (3.40), 0–15

Notes: IFIPT hits are out of 10 possible points. IFIPT = Indiana faces in places test; aMCI = amnestic mild cognitive impairment; NC = normal comparison.
*T-tests significant at $p < .05$.
**T-tests significant at $p < .01$.
***T-tests significant at $p < .001$.

Table 4. Sensitivity, specificity, positive and negative predictive powers for main IFIPT variables at baseline

	Sensitivity	Specificity	PPP	NPP
IFIPT total learning				
≤4	0.44	0.87	0.53	0.82
≤5	0.58	0.71	0.40	0.83
≤6	0.67	0.63	0.38	0.85
≤7	0.67	0.50	0.31	0.82
≤8	0.75	0.36	0.28	0.81
IFIPT delay hits				
≤3	0.57	0.70	0.38	0.84
≤4	0.74	0.44	0.30	0.84
≤5	0.86	0.33	0.29	0.88
≤6	0.91	0.18	0.26	0.86

Notes: IFIPT = Indiana faces in places test; PPP = positive predictive power; NPP = negative predictive power.
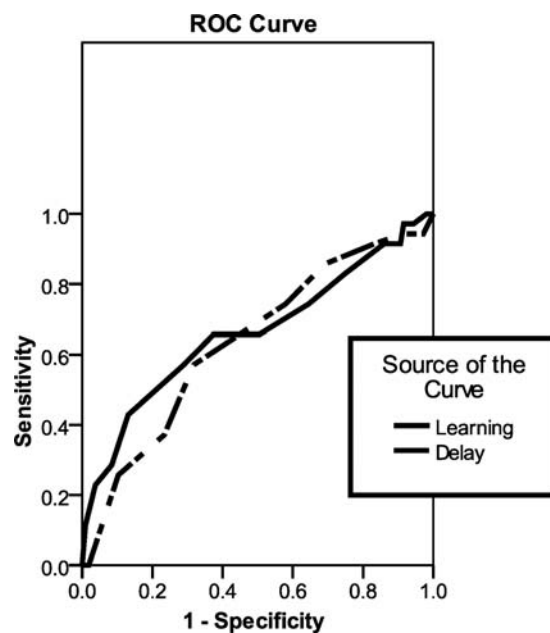
Fig. 2. Receiver operating characteristic curve for Indiana faces in places test total learning and delay hits at baseline.

0.30s to low 0.80s across at least 4 months (Levine, Miller, Becker, Selnes, & Cohen, 2004; Salinsky, Storzbach, Dodrill, & Binder, 2001; Thomas, Lawler, Olson, & Aguirre, 2008). A recently developed face perception battery (Thomas et al., 2008) also reported retest correlations in this range (0.37–0.75) over a brief retest interval of 3 weeks. The BVMT-R retest correlations presented in the manual (for the same form) range between 0.60 and 0.84 with an average interval of 56 days. Our control participants' test–retest correlations on the BVMT-R (same form) ranged from 0.67 to 0.75 over 1 week. The HVLT-R retest correlations presented in the manual for total recall and delayed recall (alternate forms) were 0.74 and 0.66, respectively, across a 6-week interval and ranged between 0.54 and 0.60 in our sample using the same form. So although a possible caveat in this study is that our first retest interval after 1 week was very short and may have inflated our retest correlations, the retest correlations for other established memory tests were lower in our sample across 1 week than those reported in their manuals over slightly longer intervals, which argues against inflation. Additionally, the IFIPT retest correlations were similar in the MCI group across 1 week and 1 year. In the control group, the 1-year values were lower than the 1-week values (raising the possibility of inflation) but within the range reported in the literature for similar tests.

There have been mixed results in the literature about whether patients with MCI demonstrate practice effects (Cooper, Lacritz, Weiner, Rosenberg, & Cullum, 2004; Duff et al., 2008). Patients with MCI are capable of learning and in some instances show greater improvement than controls on cognitive tasks (Belleville et al., 2006; Cipriani, Bianchetti, & Trabucchi, 2006; Wenisch et al., 2007). In a study using face–name pairs, patients with MCI were capable of improved

performance after training and also showed generalized improvements on other cognitive tasks. The benefit persisted for at least 1 month (Hampstead, Sathian, Moore, Nalisnick, & Stringer, 2008). Our participants in both groups showed robust practice effects across 1 week. Across 1 year, the controls showed improvement but the MCI participants showed stable performance. The fact that MCI participants did not decline across 1 year may be unexpected in a memory-impaired sample. To rule out that the lack of decline was a function of the IFIPT, we also examined delayed recall performance on the other two memory measures and found that the MCI group did not decline on the BVMT-R, and actually showed slight improvement on the HVLT-R over a year. Results are counterintuitive, but recent research has shown that fewer people with MCI go on to develop dementia than previously thought (Anstey et al., 2008; de Jager & Budge, 2005; Mitchell & Shiri-Feshki, 2009). This raises questions about the heterogeneity of etiologies contained within the MCI concept, which has implications for mixed research findings regarding cognitive changes over time. MCI remains an evolving diagnosis (Winblad et al., 2004). Additionally, the short retest interval created a "dual baseline" which may have impacted the 1-year scores. Previous literature has shown that for many tests, the majority of improvement from practice effects occurs from the first to second administration (Baird, Tombaugh, & Francis, 2007; Beglinger et al., 2005), but that some of the benefit from practice persists for months, even in an impaired group (Baird et al., 2007; Duff, Westervelt, McCaffrey, & Haase, 2001). Thus, it is possible that the lack of decline at the 1-year mark in the MCI group is a function of the dual baseline.

We provide preliminary sensitivity and specificity results for the IFIPT. The results indicated modest sensitivity, specificity, and predictive power. Although the majority of the IFIPT variables were statistically different between the MCI and control groups, the diagnostic accuracy was not high, but was consistent with other studies in which MCI is compared with controls (e.g., De Jager, Hogervorst, Combrinck, & Budge, 2003). The broad criteria used to define MCI may partially explain the fact that sensitivity and specificity are lower in MCI samples than in those with dementia. In clinical practice, it may be more feasible to tailor diagnostic decisions to the individual with some flexibility to take multiple sources of information and test data into account. For research purposes, more standardized cut-off scores are necessary, but those are somewhat arbitrary and probably result in mixed groups rather than distinct groups, which will affect diagnostic accuracy. Scores in a mildly impaired sample will not separate from controls to the same degree as they would in clearly impaired (e.g., demented) groups. Examination of the distribution of scores for the two participant groups supports this claim; the majority of both groups' participants scored in the middle of the possible score ranges. With that caveat in mind, the data in Table 4 can provide preliminary information to guide selection of cut-off scores for impairment. A cut-off of five for total learning and three for delayed hits provides a good balance between sensitivity and specificity. However, these scores should be considered starting points that require validation in a larger sample.

The IFIPT showed evidence of convergent validity with the other memory measures in our battery. In this first study of the IFIPT, we focus on the associations between the IFIPT and other measures within the control group (*see* Table 2 for correlations within the MCI group). The IFIPT correlated moderately with the other visual memory measures, the BVMT-R and the RBANS figure recall, as well as with the verbal memory measure, the HVLT-R. Although the correlations are modest with the visual memory measures in the battery, this is not altogether unexpected given that the type of visual stimuli used (i.e., faces) was different in the IFIPT compared with the geometric figures in the RBANS and BVMT-R. There is accumulating evidence that memory for faces is a special instance of visual memory, in part due to the complexity of the stimuli (Werheid & Clare, 2007). In the present study, the sample and measures were not specifically designed to validate the IFIPT. Additional research is needed to establish validity between the IFIPT and other facial memory and more general visual perceptual measures. Moderate correlations with the verbal memory measure may be a function of facial stimuli being more amenable to verbal encoding strategies than abstract visual stimuli or the reduced lateral specificity of visual memory measures. For example, the HVLT-R and BVMT-R report in their manuals that they are correlated with each other between 0.33 and 0.74. We also examined the relationship between the IFIPT and the non-memory measures for evidence of divergent validity. It was not associated with WRAT-3 reading, which is an advantage in separating premorbid level from current memory performance, nor with the COWAT or digit span. It was only associated weakly with RBANS picture naming (delay hits only). However, the IFIPT showed a surprising mild to moderate correlation with TMT, animal naming, and the SDMT. All three tasks require working memory, as does the IFIPT. In sum, the IFIPT was moderately associated with other established non-facial memory measures, but divergent validity was mixed and further validation is required.

Given the paucity of visuospatial memory measures, the current findings suggest that the IFIPT may be a useful clinical tool that requires further investigation. Although facial recognition has been dissociated from facial memory and may be subserved in part by the left hemisphere, facial learning, and memory have consistently been localized to the right temporal lobe (Barr, 1997; Broad, Mimmack, & Kendrick, 2000; Dade & Jones-Gotman, 2001; Grafman, Salazar, Weingartner, & Amin, 1986; Schiltz et al., 2006). Thus, a facial learning test has the potential to provide improved lateralizing information compared with other visual memory measures. The faces subtest from the Wechsler memory scale-III (Wechsler, 1997) is probably one of the best-known clinical measures of this kind. However, despite its relatively recent development, it has been criticized

in the literature for its lack of sensitivity (Hawkins & Tulsky, 2004; Levy, 2006; McCue, Bradshaw, & Burns, 2000) and floor effects. Additionally, as Dade and Jones-Gotman (2001) have noted, in facial memory paradigms with a single exposure to the stimuli, poor recall may be due to inattention during encoding or poor comprehension of instructions rather than a memory failure. For this reason, tests with multiple exposures to the stimuli over learning trials, such as in the IFIPT and BVMT-R, are preferred. Thus, the IFIPT has the advantages of multiple learning trials, delayed recognition trial, and it avoids ceiling effects during learning in patients with subtle memory impairment. Although the potential lateralizing effects of the IFIPT are not addressed here given the sample without focal lesions, this is a future direction for this new measure.

Some limitations of the current study should be noted. First, the current study provides only preliminary evidence for the IFIPT in a relatively small sample of MCI participants. It is unclear how these results would generalize to a larger sample, thus it is not our intention for the current results to be used as normative data for the IFIPT. It will be important to study a larger group of MCI patients, as well as to expand to patients with focal brain lesions for validation. Second, there was evidence of practice effects, particularly between the first two administrations of the test in both groups. Examination of Table 3 shows that there was improvement at all sessions for both groups. This suggests that an alternate form of the IFIPT may be useful. Third, the neuropsychological test battery was not comprehensive. There may have been deficits in cognitive domains that were not assessed by the current battery, which leaves open the possibility that some of our MCI participants might have been more impaired than they appeared. A related point concerns our selection criteria for assignment to either the MCI or NC groups. Use of an impairment score of 1.5 *SD* below average on a memory measure in the absence of notable functional decline is common for both clinical diagnosis and research in MCI, but this semi-arbitrary cut-off score may have resulted in some participant misclassification and led to heterogeneous rather than distinct groups which may have attenuated group differences and sensitivity analyses. Finally, preliminary validation of the IFIPT against two common memory measures (including one visual memory test) was provided here, but it was not validated against other measures of visuospatial perception or visual recognition memory.

In conclusion, the IFIPT is a new measure of facial learning and memory that has the benefits of being motor-free and containing multiple learning trials. In this preliminary study, the IFIPT showed moderate test–retest reliability and correlated moderately with other visual (non-facial) memory measures. It also showed clinical utility in discriminating between a sample of normal controls and participants with MCI. Additional work is needed to examine the IFIPT in other patient groups and to validate it against existing facial memory and visuospatial measures.

Conflict of Interest

None declared.

References

Alescio-Lautier, B., Michel, B. F., Herrera, C., Elahmadi, A., Chambon, C., Touzet, C., et al. (2007). Visual and visuospatial short-term memory in mild cognitive impairment and Alzheimer disease: Role of attention. *Neuropsychologia*, *45 (8)*, 1948–1960.

Anstey, K. J., Cherbuin, N., Christensen, H., Burns, R., Reglade-Meslin, C., Salim, A., et al. (2008). Follow-up of mild cognitive impairment and related disorders over four years in adults in their sixties: The PATH Through Life Study. *Dementia and Geriatric Cognitive Disorders*, *26 (3)*, 226–233.

Baird, B. J., Tombaugh, T. N., & Francis, M. (2007). The effects of practice on speed of information processing using the adjusting-paced serial addition test (adjusting-PSAT) and the computerized tests of information processing (CTIP). *Applied Neuropsychology*, *14 (2)*, 88–100.

Barr, W. B. (1997). Examining the right temporal lobe's role in nonverbal memory. *Brain and Cognition*, *35 (1)*, 26–41.

Beglinger, L. J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D. A., Crawford, J., et al. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology*, *20 (4)*, 517–529.

Belleville, S., Gilbert, B., Fontaine, F., Gagnon, L., Menard, E., & Gauthier, S. (2006). Improvement of episodic memory in persons with mild cognitive impairment and healthy older adults: Evidence from a cognitive intervention program. *Dementia and Geriatric Cognitive Disorders*, *22 (5–6)*, 486–499.

Benedict, R. H. B. (1997). *Brief visuospatial memory test-revised*. Lutz, FL: Psychological Assessment Resources.

Benton, A. L., Hamsher, K., Varney, N., & Spreen, O. (1983). *Contributions to neuropsychological assessment: A clinical manual*. New York: Oxford University Press.

Blackwell, A. D., Sahakian, B. J., Vesey, R., Semple, J. M., Robbins, T. W., & Hodges, J. R. (2004). Detecting dementia: Novel neuropsychological markers of preclinical Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *17 (1–2)*, 42–48.

Brandt, J., & Benedict, R. H. B. (2001). *Hopkins verbal learning test-revised*. Lutz, FL: Psychological Assessment Resources.

Broad, K. D., Mimmack, M. L., & Kendrick, K. M. (2000). Is right hemisphere specialization for face discrimination specific to humans? *The European Journal of Neuroscience*, *12 (2)*, 731–741.

Brown, S. A., Tapert, S. F., Granholm, E., & Delis, D. C. (2000). Neurocognitive functioning of adolescents: Effects of protracted alcohol use. *Alcoholism, Clinical and Experimental Research*, *24 (2)*, 164–171.

Butters, N., Delis, D. C., & Lucas, J. A. (1995). Clinical assessment of memory disorders in amnesia and dementia. *Annual Review of Psychology*, *46*, 493–523.

Butters, N., Wolfe, J., Martone, M., Granholm, E., & Cermak, L. S. (1985). Memory disorders associated with Huntington's disease: Verbal recall, verbal recognition and procedural memory. *Neuropsychologia*, *23 (6)*, 729–743.

Cipriani, G., Bianchetti, A., & Trabucchi, M. (2006). Outcomes of a computer-based cognitive rehabilitation program on Alzheimer's disease patients compared with those on patients affected by mild cognitive impairment. *Archives of Gerontology and Geriatrics*, *43 (3)*, 327–335.

Cooper, D. B., Lacritz, L. H., Weiner, M. F., Rosenberg, R. N., & Cullum, C. M. (2004). Category fluency in mild cognitive impairment: Reduced effect of practice in test–retest conditions. *Alzheimer Disease and Associated Disorders*, *18 (3)*, 120–122.

Dade, L. A., & Jones-Gotman, M. (2001). Face learning and memory: The twins test. *Neuropsychology*, *15 (4)*, 525–534.

Davis, H. S., & Rockwood, K. (2004). Conceptualization of mild cognitive impairment: A review. *International Journal of Geriatric Psychiatry*, *19 (4)*, 313–319.

de Jager, C. A., & Budge, M. M. (2005). Stability and predictability of the classification of mild cognitive impairment as assessed by episodic memory test performance over time. *Neurocase*, *11 (1)*, 72–79.

De Jager, C. A., Hogervorst, E., Combrinck, M., & Budge, M. M. (2003). Sensitivity and specificity of neuropsychological tests for mild cognitive impairment, vascular cognitive impairment and Alzheimer's disease. *Psychological Medicine*, *33 (6)*, 1039–1050.

Duff, K., Beglinger, L. J., Van Der Heiden, S., Moser, D. J., Arndt, S., Schultz, S. K., et al. (2008). Short-term practice effects in amnestic mild cognitive impairment: Implications for diagnosis and treatment. *International Psychogeriatrics*, *20 (5)*, 986–999.

Duff, K., Westervelt, H. J., McCaffrey, R. J., & Haase, R. F. (2001). Practice effects, test–retest stability, and dual baseline assessments with the California verbal learning test in an HIV sample. *Archives of Clinical Neuropsychology*, *16 (5)*, 461–476.

Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in the early detection of dementia. *Applied Neuropsychology*, *2 (2)*, 72–78.

Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1997). Computerized neuropsychological tests in the early detection of dementia: Prospective findings. *Journal of the International Neuropsychological Society*, *3 (2)*, 139–146.

Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (2002). Paired associate performance in the early detection of DAT. *Journal of the International Neuropsychological Society*, *8 (1)*, 58–71.

Grafman, J., Salazar, A. M., Weingartner, H., & Amin, D. (1986). Face memory and discrimination: An analysis of the persistent effects of penetrating brain wounds. *International Journal of Neuroscience*, *29 (1–2)*, 125–139.

Griffith, H. R., Netson, K. L., Harrell, L. E., Zamrini, E. Y., Brockington, J. C., & Marson, D. C. (2006). Amnestic mild cognitive impairment: Diagnostic outcomes and clinical prediction over a two-year time period. *Journal of the International Neuropsychological Society*, *12 (2)*, 166–175.

Hampstead, B. M., Sathian, K., Moore, A. B., Nalisnick, C., & Stringer, A. Y. (2008). Explicit memory training leads to improved memory for face–name pairs in patients with mild cognitive impairment: Results of a pilot investigation. *Journal of the International Neuropsychological Society*, *14 (5)*, 883–889.

Hawkins, K. A., & Tulsky, D. S. (2004). Replacement of the faces subtest by visual reproductions within Wechsler memory scale-third edition (WMS-III) visual memory indexes: Implications for discrepancy analysis. *Journal of Clinical and Experimental Neuropsychology*, *26 (4)*, 498–510.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4 (6)*, 223–233.

Jenkins, L., Myerson, J., Joerding, J. A., & Hale, S. (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. *Psychology and Aging*, *15 (1)*, 157–175.

Larrabee, G. J., & Crook, T. (1988). A computerized everyday memory battery for assessing treatment effects. *Psychopharmacology Bulletin*, *24 (4)*, 695–697.

Levine, A. J., Miller, E. N., Becker, J. T., Selnes, O. A., & Cohen, B. A. (2004). Normative data for determining significance of test–retest differences on eight common neuropsychological instruments. *The Clinical Neuropsychologist*, *18 (3)*, 373–384.

Levy, B. (2006). Increasing the power for detecting impairment in older adults with the faces subtest from Wechsler memory scale-III: An empirical trial. *Archives of Clinical Neuropsychology*, *21 (7)*, 687–692.

Lezak, M. D. (1979). Recovery of memory and learning functions following traumatic brain injury. *Cortex*, *15 (1)*, 63–72.

Lezak, M., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.

Mapstone, M., Steffenella, T. M., & Duffy, C. J. (2003). A visuospatial variant of mild cognitive impairment: Getting lost between aging and AD. *Neurology*, *60 (5)*, 802–808.

McCue, R., Bradshaw, A. A., & Burns, W. J. (2000). WMS-III subtest failure in the assessment of patients with memory impairment. *Archives in Clinical Neuropsychology*, *15*, 744.

Mitchell, A. J., & Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia—meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica*, *119 (4)*, 252–265.

Morris, R., Evenden, J., Sahakian, B., & Robbins, T. (1987). Computer-aided assessment of dementia: Comparative studies of neuropsychological deficits in Alzheimer type dementia and Parkinson's disease. In S. Stahl, S. Iversen, & E. Goodman (Eds.), *Cognitive neurochemistry* (pp. 21–36). Oxford, England: Oxford University Press.

Paulsen, J. S., Heaton, R. K., Sadek, J. R., Perry, W., Delis, D. C., Braff, D., et al. (1995). The nature of learning and memory impairments in schizophrenia. *Journal of the International Neuropsychological Society*, *1 (1)*, 88–99.

618 *L.J. Beglinger et al. / Archives of Clinical Neuropsychology 24 (2009) 607–618*

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.

Randolph, C. (Ed.). (1998). *Repeatable battery for the assessment of neuropsychological status (manual)*. San Antonio, TX: The Psychological Corporation.

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*, 271–276.

Salinsky, M. C., Storzbach, D., Dodrill, C. B., & Binder, L. M. (2001). Test–retest bias, reliability, and regression equations for neuropsychological measures repeated over a 12–16-week period. *Journal of the International Neuropsychological Society*, *7 (5)*, 597–605.

Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., et al. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cerebral Cortex*, *16 (4)*, 574–586.

Smith, A. (1991). *Symbol digits modalities test*. Los Angeles: Western Psychological Services.

Teng, E. L., & Chui, H. C. (1987). The modified mini-mental state (3MS) examination. *Journal of Clinical Psychiatry*, *48 (8)*, 314–318.

Thomas, A. L., Lawler, K., Olson, I. R., & Aguirre, G. K. (2008). The Philadelphia face perception battery. *Archives of Clinical Neuropsychology*, *23 (2)*, 175–187.

Wechsler, D. (1997). *Wechsler memory scale—Third edition*. San Antonio, TX: The Psychological Corporation.

Wenisch, E., Cantegreil-Kallen, I., De Rotrou, J., Garrigue, P., Moulin, F., Batouche, F., et al. (2007). Cognitive stimulation intervention for elders with mild cognitive impairment compared with normal aged subjects: Preliminary results. *Aging Clinical and Experimental Research*, *19 (4)*, 316–322.

Werheid, K., & Clare, L. (2007). Are faces special in Alzheimer's disease? Cognitive conceptualisation, neural correlates, and diagnostic relevance of impaired memory for faces and names. *Cortex*, *43 (7)*, 898–906.

Wilkinson, G. S. (1993). *Wide range achievement test—Third edition*. Wilmington, DE: Wide Range Incorporated.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., et al. (2004). Mild cognitive impairment–beyond controversies, towards a consensus: Report of the International Working Group on mild cognitive impairment. *Journal of Internal Medicine*, *256 (3)*, 240–246.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*, 37–49.