Published in final edited form as:
Int Psychogeriatr. 2008 October ; 20(5): 986–999. doi:10.1017/S1041610208007254.

Short-term practice effects in amnestic mild cognitive impairment: implications for diagnosis and treatment

Kevin Duff[a1,c1], **Leigh J. Beglinger**[a1], **Sara Van Der Heiden**[a1], **David J. Moser**[a1], **Stephan Arndt**[a1,a2], **Susan K. Schultz**[a1], and **Jane S. Paulsen**[a1]

[a1]Department of Psychiatry, College of Medicine, University of Iowa, Iowa City, U.S.A

[a2]Department of Biostatistics, College of Public Health, University of Iowa, Iowa City, U.S.A

Abstract

Background—Practice effects have been widely reported in healthy older adults, but these improvements due to repeat exposure to test materials have been more equivocal in individuals with mild cognitive impairment (MCI).

Methods—The current study examined short-term practice effects in MCI by repeating a brief battery of cognitive tests across one week in 59 older adults with amnestic MCI and 62 intact older adults.

Results—Participants with amnestic MCI showed significantly greater improvements on two delayed recall measures ($p < 0.01$) compared to intact peers. All other practice effects were comparable between these two groups. Practice effects significantly improved scores in the MCI group so that 49% of them were reclassified as “intact” after one week, whereas the other 51% remained “stable” as MCI. Secondary analyses indicated the MCI-Intact group demonstrated larger practice effects on two memory measures than their peers ($p < 0.01$).

Conclusions—These results continue to inform us about the nature of memory deficits in MCI, and could have implications for the diagnosis and possible treatment of this amnestic condition.

Keywords

mild cognitive impairment; practice effects; repeat testing

Introduction

Practice effects are defined as improvements in cognitive test performance due to repeated evaluation with the same test materials, and have traditionally been viewed as sources of error (McCaffrey *et al.*, 2000). Whereas practice effects have been widely reported in cognitively intact older adults (McCaffrey *et al.*, 2000; Beglinger *et al.*, 2005b) and largely absent in patients with dementia (Cooper *et al.*, 2001; Helkala *et al.*, 2002), less is known about practice effects in amnestic Mild Cognitive Impairment (MCI). Darby *et al.* (2002)

[c1]Correspondence should be addressed to: Dr. Kevin Duff, University of Iowa, Department of Psychiatry, MEB 1-308, Iowa City, IA 52242-1000, U.S.A. Phone: +1 319 335 6640; Fax: +1 319 353 3003. kevin-duff@uiowa.edu.

Conflict of interest: None.

Description of authors' roles

K. Duff and L. Beglinger were responsible for study concept, data acquisition, analysis and interpretation of the data, and the preparation of the manuscript. S. Van Der Heiden was primarily responsible for data acquisition and processing of the data. D. Moser, S. Schultz, and J. Paulsen assisted in developing the study concept and design and interpretation of the data. S. Arndt assisted in the study design, planned the data analysis, and assisted in the interpretation of the data. All authors reviewed the manuscript and approved the final version.

reported an absence of practice effects on a computerized battery of cognitive tasks repeated within a single day in patients with MCI. Over a slightly longer retest period (i.e. one week), practice effects were also absent in a group of patients with MCI on a semantic fluency task (Cooper *et al.*, 2004). Similarly, Galvin *et al.* (2005) and others (Schrijnemaekers *et al.*, 2006) have reported an absence of practice effects over much longer periods (i.e. 1–3 years) in those progressing from MCI to dementia.

Conversely, individuals with amnestic MCI have been reported to demonstrate practice effects on cognitive and motor tests across brief retest periods (Yan and Dick, 2006; Duff *et al.*, 2007). The equivocal findings in the literature could be attributed to several factors, including small sample sizes (Duff *et al.*, 2007), limited assessment batteries (Cooper *et al.*, 2004; Schrijnemaekers *et al.*, 2006; Yan and Dick, 2006), or long retest periods (Galvin *et al.*, 2005; Schrijnemaekers *et al.*, 2006). Given these discrepancies in the literature, the current study sought to compare practice effects in older adults with amnestic MCI with peers with intact cognition. Although the literature in this area is mixed, it is expected that individuals with primary memory deficits (i.e. amnestic MCI) would display smaller practice effects than individuals with intact cognition.

In addition to learning more about memory abilities in this amnestic condition, practice effects data in MCI could be useful in other ways. First, such data could inform the design and interpretation of drug trials for this prodromal phase of dementia (Beglinger *et al.*, 2005a). For example, it is necessary to know whether some degree of practice effects are expected in amnestic MCI participants before initiation of a drug. Second, practice effects could provide clinically useful information about progression of the illness (Duff *et al.*, 2007). Third, it is likely that some tests are more susceptible to practice effects, whereas others are not, and a more comprehensive examination of practice effects across multiple cognitive domains could be fruitful. Finally, it is possible that certain patient characteristics (e.g. age, education, baseline cognitive functioning) lead to differential practice effects (McCaffrey *et al.*, 2000).

Methods

Participants and procedures

All procedures were approved by the local Institutional Review Board. One hundred and twenty-one older adults were recruited from independent living facilities and community senior centers following educational talks on cognitive changes associated with aging. All participants denied having a history of major neurological (e.g. traumatic brain injury, stroke, dementia) or psychiatric illness (e.g. schizophrenia, bipolar disorder) or current depression (either self-report or 30-item Geriatric Depression Scale (GDS) score of >12). All participants completed a brief telephone screening (Lines *et al.*, 2003), which has been shown to assist in identifying amnestic MCI. All of these individuals completed a baseline assessment, which included: clinical interview, GDS, Wide Range Achievement Test–3 (WRAT-3) Reading subtest, modified Mini-mental State Examination (3MS) temporal and spatial orientation items, Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), Brief Visuospatial Memory Test – Revised (BVMT-R), Hopkins Verbal Learning Test – Revised (HVLT-R), Controlled Oral Word Association Test (COWAT), animal fluency, Trail Making Test (TMT) parts A and B, and Symbol Digit Modalities Test (SDMT). Using results from the baseline assessment, individuals were classified as cognitively intact or amnestic MCI using existing criteria (Petersen *et al.*, 1999). To be classified as amnestic MCI, all participants had to complain of memory problems (self-reported as yes/no during an interview). These participants had objective memory deficits (i.e. age-corrected scores at or below the 7th percentile on at least two of the three delayed recall measures (RBANS Delayed Memory Index, HVLT-R, BVMT-R) relative to a

premorbid intellectual estimate (WRAT-3 Reading)). The 7th percentile is 1.5 standard deviations below the mean, which is a typical demarcation point for cognitive deficits in MCI. Cognition was otherwise generally intact (i.e. non-memory age-corrected scores above the 7th percentile) and no functional impairments (e.g. assistance needed with managing money, taking medications, driving) could be reported. To be classified as "cognitively intact," all objective memory and non-memory performances were at least above the 7th percentile. All data were reviewed by two neuropsychologists (KD, LJB), and individuals were classified with amnestic MCI (n = 59) or cognitively intact (n = 62). No one was classified as demented (i.e. with both impaired memory and other cognitive domains). All classifications were made following the one-week visit, so examiners were "blinded" to classification at the baseline and one-week visits. However, only baseline cognitive performances were used in these classifications. Demographic and baseline assessment scores are presented in Tables 1 and 2.

Approximately one week after the baseline visit, all participants completed a repeat cognitive assessment, which included all the baseline measures except WRAT-3 Reading, 3MS, and RBANS. Since practice effects were the primary focus of this study, alternate forms of the tests were not used to maximize practice effects.

Data analysis

Practice effects were calculated for all nine repeated tests as one-week score/baseline score. Raw scores were used for all practice effects scores, except the BVMT-R and HVLT-R. Age-corrected standard scores from their respective manuals were used for the total learning and delayed recall scores of the BVMT-R and HVLT-R to avoid zeros in either the numerator or denominator of our practice effects scores. Additionally, scores on the TMT were reversed so that lower scores indicated poorer performances, which was then consistent with all the other measures. These individual practice effects scores are ratios of one-week follow-up to baseline scores, with 1.0 indicating no change, >1.0 indicating improvement on follow-up, and <1.0 indicating decline on follow-up. For these individual practice effects scores, a score of 1.2 would indicate that the one-week score was 120% of the baseline score or an improvement of 20% from baseline. Conversely, a score of 0.8 would indicate that the one-week score was 80% of the baseline score or a decline of 20% from baseline.

Demographics, WRAT-3 standard scores, and GDS raw scores were compared with independent t-tests and χ^2 analyses. As noted below, the two groups were significantly different on age ($p < 0.01$), so age was used as a covariate in all the remaining analyses. Two MANCOVAs were used to examine baseline differences between the groups on the cognitive measures. The first MANCOVA (controlling for age) compared all baseline non-memory measures (i.e. 3MS, visuospatial/constructional, language, and attention indexes of the RBANS, COWAT, animal fluency, TMT, SDMT). The second MANCOVA (also controlling for age) compared the groups on memory measures (i.e. immediate and delayed memory indexes of the RBANS, total and delayed recall of the HVLT-R, total and delayed recall of the BVMT-R). If these two groups actually represented amnestic MCI and intact cognition, then there would be no differences on the first MANCOVA (non-memory tasks) but significant differences on the second one (memory tasks). These two MANCOVAs were validity checks of our classification method and not the primary hypothesis of interest.

The primary outcome measures, the individual practice effects scores, were compared with a MANCOVA, controlling for age. The alpha level was set at 0.01 to decrease the risk of a Type I error due to multiple comparisons in the post-hoc analyses.

Results

MCI classification based on baseline assessment

The amnestic MCI group was significantly older than the intact group ($F[1,120] = 16.0$, $p < 0.001$), but the groups were comparable for education ($p = 0.93$), gender ($p = 0.79$), estimated premorbid intellect ($p = 0.07$) and depression ($p = 0.06$). All participants in both groups were Caucasian. After controlling for age, the groups were comparable on all non-memory tests at the baseline assessment (multivariate $F[9,106] = 1.20$, $p = 0.30$). Consistent with existing criteria (Petersen *et al.*, 1999), the amnestic MCI group performed significantly below their healthy peers on all tests of immediate and delayed memory (multivariate $F[6,113] = 23.1$, $p < 0.001$, partial $\eta^2 = 0.55$). The results of these two MANCOVAs support the classification of participants as amnestic MCI or intact.

Primary analyses of practice effects between groups

The amnestic MCI and intact groups did not differ in retest interval (MCI = 7.6 [3.1] days; intact = 7.4 [1.6] days, $p = 0.52$). The MANCOVA on all nine practice effects scores indicated a significant group effect (multivariate $F[9,104] = 3.8$, $p < 0.001$, partial $\eta^2 = 0.25$), with the largest improvements for the MCI participants on the BVMT-R delayed recall ($p < 0.001$, partial $\eta^2 = 0.13$) and HVLT-R delayed recall ($p = 0.002$, partial $\eta^2 = 0.08$). Practice effects for the other repeated tests were comparable between the two groups ($p > 0.01$). Practice effects ratio scores are presented in Table 3.

MCI reclassification based on one-week assessment

Since a number of the participants originally classified as amnestic MCI significantly improved on the cognitive testing between the baseline and one-week assessments, we attempted to reclassify participants based on one-week performances as either amnestic MCI or intact using procedures similar to those described above for the original classification of participants. Briefly, two neuropsychologist (KD, LJB) reviewed all one-week data and reclassified participants as amnestic MCI or intact based on the presence or absence of objective memory deficits on both of the delayed recall measures from the one-week assessment (HVLT-R, BVMT-R) relative to a premorbid intellectual estimate. Results of this reclassification yielded 35 amnestic MCI participants and 86 cognitively intact participants. Over the course of one week, no participant significantly declined in non-memory performances to be classified as demented. All reclassifications were made independent of and “blinded” to the original classifications.

Since 49% of participants originally classified as amnestic MCI reverted to intact at reclassification, it was decided to place all participants into one of three groups: (1) MCI Stable (i.e. originally classified and reclassified after one week as amnestic MCI, n = 30); (2) MCI-Intact (i.e. originally classified as amnestic MCI but reclassified after one week as intact, n = 29); and (3) Intact Stable (i.e. originally classified and reclassified after one week as intact, n = 57). Five individuals originally classified as intact were reclassified as MCI at one week, but this group was considered too small to include in the remaining analyses of practice effects. Demographic and baseline and one-week assessment scores for these three groups are presented in Table 4 and 5.

The MCI Stable and MCI-Intact groups were significantly older than the Intact Stable group ($F[2,115] = 6.7$, $p < 0.001$), but the groups were comparable for education ($p = 0.17$), gender ($p = 0.55$), estimated premorbid intellect ($p = 0.23$) and baseline depression ($p = 0.19$). Unlike the initial analyses comparing MCI and Intact, after controlling for age, there were significant differences among the three groups on non-memory tests at the baseline assessment (multivariate $F[24,196] = 2.7$, $p<0.001$, partial $\eta^2 = 0.25$). Post-hoc analyses

indicated that group differences occurred on two RBANS subtests (semantic fluency p = 0.003, partial $\eta^2 = 0.10$), coding ($p < 0.001$), partial $\eta^2 = 0.20$), animal fluency (p = 0.004, partial $\eta^2 = 0.10$), Trail Making Test Part B ($p < 0.001$, partial $\eta^2 = 0.19$), and SDMT ($p < 0.001$, partial $\eta^2 = 0.15$), with the MCI Stable group consistently performing below the Intact Stable group, and the MCI-Intact group falling between the other two groups. Similar to our original classification results, baseline memory tests were significantly different among the three groups (multivariate $F[12,216] = 9.5$, $p < 0.001$, partial $\eta^2 = 0.35$). Post hoc analyses revealed that group differences occurred on all immediate and delayed memory measures (all $p < 0.001$, partial η^2 range: 0.18–0.50, MCI Stable < MCI-Intact < Intact Stable).

Secondary analyses of practice effects between groups

The three groups did not differ in retest interval (p = 0.34). The MANCOVA on all nine practice effects scores indicated a significant group effect (multivariate $F[18,198] = 2.9$, $p < 0.001$, partial $\eta^2 = 0.21$), after controlling for age. Post-hoc analyses indicated that significant group differences occurred on the BVMT-R delayed recall ($p < 0.001$, partial $\eta^2 = 0.13$, MCI Stable = MCI-Intact > Intact Stable), HVLT-R total recall (p = 0.002, partial $\eta^2 = 0.11$, MCI-Intact > Intact Stable = MCI Stable), and HVLT-R delayed recall (p = 0.009, partial $\eta^2 = 0.08$, MCI Stable > Intact Stable). Practice effects for the other repeated tests were comparable between the three groups ($p > 0.01$). Practice effects ratio scores are presented in Table 6.

Discussion

In the primary analyses, both intact and amnestic MCI participants displayed practice effects across one week, with improved test performances on a variety of cognitive measures. Individuals identified as intact at the baseline visit demonstrated the expected practice effects on all measures across this brief retest interval (Benedict and Zgaljardic, 1998; Bird *et al.*, 2004; Beglinger *et al.*, 2005b), with gains of 6–30% across one week. Unexpectedly, however, the individuals classified as amnestic MCI at baseline improved significantly more than their intact peers on two memory measures (BVMT-R delayed recall: 32% vs. 13% improvement for MCI vs. intact, respectively; and HVLT-R delayed recall: 26% vs. 9% improvement). These findings conflict with some existing literature in patients with early dementia and MCI (Darby *et al.*, 2002; Cooper *et al.*, 2004; Galvin *et al.*, 2005; Schrijnemaekers *et al.*, 2006), where practice effects have been largely absent. Some of the discrepancies between our findings and other studies in the literature are expected. For example, two of these prior studies (Galvin *et al.*, 2005; Schrijnemaekers *et al.*, 2006) used very long retest intervals (e.g. 1–3 years) and practice effects are likely to be attenuated across such periods. Additionally, other studies have used alternate test forms to minimize practice effects, whereas our procedures tried to maximize practice effects by purposely avoiding alternate forms. Our current results do converge with our previous, but independent, smaller study (Duff *et al.*, 2007), which found increased practice effects in amnestic MCI across similar retest intervals and without alternate forms.

Given this somewhat unexpected finding (i.e. patients with amnestic MCI benefit from repeated exposure to test materials more than intact peers), it is worth discussing some possible explanations. First, from a methodological standpoint, intact individuals were more susceptible to ceiling effects than MCI participants. Since the intact participants had higher baseline scores on all memory measures than those in the MCI group, they had less room to improve on retesting, which could have diminished their practice effects on these measures. Future studies might include supra-span memory tests that minimize this potential confound. Second, from a conceptual standpoint, practice effects may tap into two memory subsystems: direct, declarative, content learning (e.g. remembering the actual words on list

learning task) and indirect, procedural, contextual learning (e.g. remembering how to solve a specific type of problem). It is possible that cognitive declines associated with MCI do not affect these two subsystems contemporaneously, and procedural learning and memory could be retained longer into the course of the illness (Yan and Dick, 2006). Along these same lines, it is possible that memory deficits in MCI are still different from those in Alzheimer's disease. For example, recent research has suggested that some patients with MCI demonstrate relatively better performance on recognition trials than recall trials (Bennett *et al.*, 2006; Westerberg *et al.*, 2006), and better recognition memory might facilitate the expression of practice effects. Lastly, from both a methodological and conceptual standpoint, amnestic MCI is typically viewed as a heterogeneous group, comprised of patients who will progress to dementia and those who will remain stable for several years or even revert back to normal cognition (Winblad *et al.*, 2004). The heterogeneity of the MCI group might be indicated by the variability in practice effects in this group, which could suggest two subgroups of patients with MCI: those who benefit from practice and those who do not. Our secondary data analyses provide further support for these subgroups of MCI. Using only one-week data, nearly half of our MCI participants improved so much that they were reclassified as "intact" after one week, whereas the other half retained their MCI status. We have previously suggested that these two subgroups might have differential courses and outcomes (Duff *et al.*, 2007), but this needs further investigation.

Regardless of the explanation for these findings, the current results appear to have clinical implications for the diagnosis and treatment of MCI. Using only baseline test data, the current memory impaired sample appears to meet criteria for amnestic MCI, with subjective complaints, significant memory deficits (e.g. mean delayed recall scores on BVMT-R and HVLT-R = 4th percentile), but relatively intact cognition (e.g. mean RBANS total score = 39th percentile). One week later, repeat testing showed dramatic improvements in memory (e.g. mean delayed recall scores on BVMT-R and HVLT-R = 42nd percentile). Clear classification of these subjects as amnestic MCI at this point would be difficult. As noted above, using only repeat (i.e. one-week follow-up) testing scores, 51% of those originally classified as MCI would retain their classification, with the remainder shifting to the intact group. Five of the intact participants shifted from their group to MCI using one-week test scores. It should be noted, however, that only two of the three memory tests were available at one-week follow-up to classify participants, which could limit the accuracy of these classifications. Additionally, to our knowledge, short-term repeat testing normative data do not exist, which makes the reclassification and secondary analyses preliminary. Nonetheless, future studies might examine the clinical utility of repeat testing for the diagnosis of amnestic MCI, as significant improvements in memory functioning (as evident with practice effects) might seriously question the validity of the original diagnosis. Furthermore, Darby *et al.* (2002) have provided data to suggest that multiple assessments within the same day could be used to identify MCI, which might be better at identifying persistent vs. "accidental" MCI (de Rotrou *et al.*, 2005).

The current findings could also have treatment implications. Even though the effect sizes for the delayed recall measures were moderate, it is informative to note that both groups were comparable on all other measures of practice effects. This could be interpreted as indicating that these individuals with amnestic MCI benefited as much from repetition as their intact peers. Even those individuals reclassified as MCI Stable after one week demonstrated improvements on retesting compared to the Intact Stable group on some measures (e.g. BVMT-R delayed recall: MCI Stable = 24% improved vs. Intact Stable = 14% improved; HVLT-R delayed recall: MCI Stable = 22% improved vs. Intact Stable = 9% improved), but not others (e.g. COWAT: MCI Stable = 1% improved vs. Intact Stable = 6% improved; SDMT: MCI Stable = 3% improved vs. Intact Stable = 9% improved). These findings might be used to guide the development of interventions for MCI (e.g. focusing on delayed recall

abilities rather than executive abilities). Overall, if some individuals with amnestic MCI benefit from practice, repetition or additional learning trials, then cognitive rehabilitation might be indicated for these patients. A few studies have found that patients with MCI benefit from cognitive interventions (Rapp *et al.*, 2002; Belleville *et al.*, 2006; Wenisch *et al.*, 2007). Additionally, cholinesterase inhibitors and other cognitive enhancing medications might work optimally in those patients who demonstrate the capacity to learn. Finally, cognitive plasticity, which might be quantified via practice effects, has been shown to be a modulating variable in the response to memory training programs in healthy elders (Calero and Navarro, 2007). Future intervention trials might consider practice effects as a variable of interest for enriching samples for intervention trials by including both groups of patients who demonstrate practice effects on short-term retesting and those that do not.

Several limitations of the current study should be noted. The participants were a high functioning group of Caucasian retirees, with an average premorbid IQ of 108. The generalizability of these findings to other samples (e.g. lower education, non-Caucasian) is unclear. Current participants did not undergo extensive medical work-ups (e.g. physical exam, neuroimaging) to confirm status, and information beyond cognitive test scores needs to be considered in participant classification. Regression to the mean could explain some of the changes in test scores. However, when we reanalyzed our data using repeated measures ANCOVAs (Barnett *et al.*, 2005), we found essentially the same findings. Our method for calculating change across time and practice effects is not the only method (e.g. subtraction method, reliable change indices, regression-based formulas), and future investigations might explore whether other methods are more sensitive at detecting change (for a review of these techniques, see Collie *et al.*, 2002). Finally, it should be reiterated that ceiling effects potentially minimized practice effects in our Intact group, and careful selection of cognitive tests in studies employing repeated assessments is warranted.

In conclusion, practice effects, frequently considered to be an error that needs to be minimized, might hold valuable information for clinicians and researchers about diagnosis and treatment in amnestic MCI. At least some patients with this amnestic condition can improve with repeated exposure to testing materials, and these results pose challenges to the definition/diagnosis of MCI but also offer hope for intervention. Practice effects, as a simple, convenient, and non-invasive marker for monitoring an individual patient's cognitive status, might also be used to offer interventions to patients who are in the earliest stages of progressive neurodegenerative disorders (e.g. enriching samples in clinical trials). We are continuing to follow this cohort, and hope to validate our previous finding that practice effects can serve as a prognostic index of future cognitive functioning.

Acknowledgments

This project was supported by a research grant (NIA R03 AG025850-01) from the National Institutes on Aging. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

References

Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. International Journal of Epidemiology. 2005; 34:215–220. [PubMed: 15333621]

Beglinger LJ, et al. Practice effects and the use of alternate forms in serial neuropsychological testing. Archives of Clinical Neuropsychology. 2005a; 20:517–529. [PubMed: 15896564]

Beglinger LJ, Tangphao-Daniels O, Kareken DA, Zhang L, Mohs R, Siemers ER. Neuropsychological test performance in healthy elderly volunteers before and after donepezil administration: a randomized, controlled study. Journal of Clinical Psychopharmacology. 2005b; 25:159–165. [PubMed: 15738747]

Belleville S, Gilbert B, Fontaine F, Gagnon L, Menard E, Gauthier S. 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. 2006; 22:486–499. [PubMed: 17050952]

Benedict RH, Zgaljardic DJ. Practice effects during repeated administrations of memory tests with and without alternate forms. Journal of Clinical and Experimental Neuropsychology. 1998; 20:339–352. [PubMed: 9845161]

Bennett IJ, Golob EJ, Parker ES, Starr A. Memory evaluation in mild cognitive impairment using recall and recognition tests. Journal of Clinical and Experimental Neuropsychology. 2006; 28:1408–1422. [PubMed: 17050267]

Bird CM, Papadopoulou K, Ricciardelli P, Rossor MN, Cipolotti L. Monitoring cognitive changes: psychometric properties of six cognitive tests. British Journal of Clinical Psychology. 2004; 43:197–210. [PubMed: 15169618]

Calero MD, Navarro E. Cognitive plasticity as a modulating variable on the effects of memory training in elderly persons. Archives of Clinical Neuropsychology. 2007; 22:63–72. [PubMed: 17158023]

Collie A, Darby DG, Falleti MG, Silbert BS, Maruff P. Determining the extent of cognitive change after coronary surgery: a review of statistical procedures. Annals of Thoracic Surgery. 2002; 73:2005–2011. [PubMed: 12078822]

Cooper DB, et al. Effects of practice on category fluency in Alzheimer's disease. The Clinical Neuropsychologist. 2001; 15:125–128. [PubMed: 11778573]

Cooper DB, Lacritz LH, Weiner MF, Rosenberg RN, Cullum CM. Category fluency in mild cognitive impairment: reduced effect of practice in test-retest conditions. Alzheimer Disease and Associated Disorders. 2004; 18:120–122. [PubMed: 15494616]

Darby D, Maruff P, Collie A, McStephen M. Mild cognitive impairment can be detected by multiple assessments in a single day. Neurology. 2002; 59:1042–1046. [PubMed: 12370459]

de Rotrou J, Wenisch E, Chausson C, Dray F, Faucounau V, Rigaud AS. Accidental MCI in healthy subjects: a prospective longitudinal study. European Journal of Neurology. 2005; 12:879–885. [PubMed: 16241977]

Duff K, et al. Practice effects in the prediction of long-term cognitive outcome in three patient samples: a novel prognostic index. Archives of Clinical Neuropsychology. 2007; 22:15–24. [PubMed: 17142007]

Galvin JE, et al. Predictors of preclinical Alzheimer disease and dementia: a clinicopathologic study. Archives of Neurology. 2005; 62:758–765. [PubMed: 15883263]

Helkala EL, et al. Usefulness of repeated presentation of Mini-mental State Examination as a diagnostic procedure – a population-based study. Acta Neurologica Scandinavica. 2002; 106:341–346. [PubMed: 12460138]

Lines CR, McCarroll KA, Lipton RB, Block GA. Telephone screening for amnestic mild cognitive impairment. Neurology. 2003; 60:261–266. [PubMed: 12552041]

McCaffrey, RJ.; Duff, K.; Westervelt, HJ. Practitioner's Guide to Evaluating Change with Neuropsychological Assessment Instruments. New York: Plenum/Kluwer; 2000.

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Archives of Neurology. 1999; 56:303–308. [PubMed: 10190820]

Rapp S, Brenes G, Marsh AP. Memory enhancement training for older adults with mild cognitive impairment: a preliminary study. Aging and Mental Health. 2002; 6:5–11. [PubMed: 11827617]

Schrijnemaekers AM, de Jager CA, Hogervorst E, Budge MM. Cases with mild cognitive impairment and Alzheimer's disease fail to benefit from repeated exposure to episodic memory tests as compared with controls. Journal of Clinical and Experimental Neuropsychology. 2006; 28:438–455. [PubMed: 16618630]

Wenisch E, et al. Cognitive stimulation intervention for elders with mild cognitive impairment compared with normal aged subjects: preliminary results. Aging: Clinical and Experimental Research. 2007; 19:316–322. [PubMed: 17726363]

Westerberg CE, et al. When memory does not fail: familiarity-based recognition in mild cognitive impairment and Alzheimer's disease. Neuropsychology. 2006; 20:193–205. [PubMed: 16594780]

Winblad B, et al. Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. Journal of Internal Medicine. 2004; 256:240–246. [PubMed: 15324367]

Yan JH, Dick MB. Practice effects on motor control in healthy seniors and patients with mild cognitive impairment and Alzheimer's disease. Aging, Neuropsychology, and Cognition. 2006; 13:385–410.

Table 1

Demographic and other baseline data in participants originally classified as amnestic MCI or intact.

DEMOGRAPHIC AND OTHER BASELTNE DATA	AMNESTIC MCI	INTACT
Age (years)	82.4 (6.4)	77.2 (7.8)*
Gender	79% female	81% female
Education (years)	15.3 (2.8)	15.4 (2.6)
WRAT-3 Reading (standard score)	109.3 (5.2)	107.6 (6.0)
GDS (raw score, max. = 30)	4.6 (3.2)	3.4 (3.3)
3MS orientation (raw score, max. = 20)	19.8 (0.4)	19.9 (0.3)
RBANS		
Immediate memory	91.1 (14.1)	105.0 (14.5)
Visuospatial constructional	105.2 (16.2)	106.6 (14.3)
Language	98.2 (10.8)	102.6 (10.7)
Attention	102.4 (12.1)	103.7 (15.0)
Delayed memory	91.0 (14.8)	106.2 (9.4)
Total	96.2 (11.2)	106.6 (11.1)

WRAT-3 = Wide Range Achievement Test-3; GDS = Geriatric Depression Scale; 3MS = Modified Mini-mental State Examination; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status.

*$p < 0.01$.

Table 2

Baseline and one-week cognitive test scores in participants originally classified as amnestic MCI or intact.

	AMNESTIC MCI				INTACT			
MEASURE	**BASELINE**	**ONE-WEEK**	**DIFFERENCE**	**r**	**BASELINE**	**ONE-WEEK**	**DIFFERENCE**	**r**
BVMT-R								
Total recall	73.1 (14.1)	91.2 (21.4)	−18.0 (15.4)	0.70	90.6 (15.6)	116.0 (16.6)	−25.4 (14.2)	0.61
Delayed recall	69.6 (15.8)	90.2 (18.4)	−20.6 (14.7)	0.64	97.7 (14.9)	109.6 (14.2)	−11.9 (11.7)	0.67
HVLT-R								
Total recall	89.9 (13.5)	103.4 (16.4)	−13.5 (11.6)	0.71	107.2 (11.9)	120.0(12.0)	−12.8 (10.7)	0.60
Delayed recall	78.4 (16.1)	97.2 (16.9)	−18.8 (14.8)	0.60	102.3 (12.6)	110.3 (8.0)	−8.1 (10.4)	0.56
COWAT	39.0 (11.7)	39.0 (11.8)	0.0 (6.9)	0.83	38.1 (10.7)	40.1 (12.8)	−2.0 (7.8)	0.79
Animals	15.6 (4.6)	15.6 (5.1)	0.0 (3.8)	0.70	18.8 (5.6)	19.4 (5.2)	−0.6 (4.4)	0.67
TMT-A	48.2 (17.5)	43.9 (14.8)	4.3 (11.6)	0.76	41.7 (13.5)	37.3 (10.6)	4.4 (8.3)	0.79
TMT-B	134.0 (63.1)	122.1 (65.2)	11.9 (46.7)	0.73	105.3 (46.3)	93.3 (35.1)	11.9 (33.8)	0.69
SDMT	36.6 (9.8)	37.5 (10.3)	−0.9 (3.9)	0.93	40.6 (8.0)	44.1 (8.7)	−3.5 (5.7)	0.77

BVMT-R = Brief Visuospatial Memory Test – Revised; HVLT-R = Hopkins Verbal Learning Test – Revised; COWAT = Controlled Oral Word Association Test; TMT = Trail Making Test; SDMT = Symbol Digit Modalities Test. Raw scores are presented for each measure, except BVMT-R and HVLT-R, which are age-corrected standard scores (M = 100, SD = 15). TMT scores are not reversed, as described in the Methods section in the calculation of practice effects scores. Difference = baseline – one-week score; r = Pearson correlations between baseline and one-week scores; r values are all $p < 0.01$.

Int Psychogeriatr. Author manuscript; available in PMC 2013 October 03.

Table 3

Ratios of practice effects in participants originally classified as amnestic MCI or intact.

PRACTICE EFFECTS	AMNESTIC MCI	INTACT
BVMT-R		
Total Recall	1.26 (0.24)	1.30 (0.20)
Delayed Recall	1.32 (0.28)	1.13 (0.14)*
HVLT-R		
Total Recall	1.16 (0.14)	1.12 (0.11)
Delayed Recall	1.26 (0.25)	1.09 (0.12)*
COWAT	1.02 (0.22)	1.06 (0.20)
Animals	1.02 (0.26)	1.08 (0.31)
TMT-A	1.05 (0.24)	1.07 (0.18)
TMT-B	1.07 (0.28)	1.06 (0.25)
SDMT	1.03 (0.11)	1.10 (0.16)

Practice effects were calculated as: one-week raw score/baseline raw score, except BVMT-R and HVLT-R, which are age-corrected standard scores (M = 100, SD = 15). BVMT-R = Brief Visuospatial Memory Test – Revised; HVLT-R = Hopkins Verbal Learning Test – Revised; COWAT = Controlled Oral Word Association Test; TMT = Trail Making Test; SDMT = Symbol Digit Modalities Test.

* $p < 0.01$ after controlling for age differences.

Int Psychogeriatr. Author manuscript; available in PMC 2013 October 03.

Table 4

Demographic and other baseline data in participants reclassified after one week as MCI Stable, MCI – Intact, or Intact Stable.

DEMOGRAPHIC AND OTHER BASELINE DATA	MCI STABLE	MCI-INTACT	INTACT STABLE
Age (years)	83.4 (6.5)	81.4 (6.5)	76.7 (7.9)
Gender	73% female	83% female	82% female
Education (years)	14.8 (2.2)	16.1 (3.3)	15.5 (2.6)
WRAT-3 Reading (standard score)	110.0 (4.9)	108.9 (5.3)	107.8 (6.1)
GDS (raw score, max. = 30)	4.9 (3.5)	4.3 (3.1)	3.6 (3.4)
3MS Orientation (raw score, max. = 20)	19.8 (0.5)	19.9 (0.4)	19.9 (0.3)
RBANS			
Immediate memory	91.0 (13.2)	93.5 (12.4)	105.2 (14.8)
Visuospatial constructional	106.0 (16.3)	105.1 (16.7)	107.2 (13.9)
Language	96.6 (10.1)	99.7 (10.3)	102.6 (11.1)
Attention	97.5 (10.5)	108.1 (11.4)	103.1 (15.0)
Delayed memory	87.7 (16.3)	96.5 (9.0)	106.5 (9.7)
Total	93.8 (11.2)	100.1 (8.8)	106.8 (11.5)

WRAT-3 = Wide Range Achievement Test – 3; GDS = Geriatric Depression Scale; 3MS = Modified Mini Mental Status Examination; RBANS = Repeatable Battery for the Assessment of Neuropsychological Status. *$p < 0.01$.

Table 5

Baseline and one-week cognitive test scores in participants reclassified after one week as MCI Stable, MCI-Intact, or Stable Intact.

	MCI STABLE		MCI-INTACT	
MEASURES	**BASELINE**	**ONE-WEEK**	**BASELINE**	**ONE-WEEK**
BVMT-R				
Total recall	67.4 (9.9)	76.5 (15.3)	79.0 (15.4)	106.4 (15.5)
Delayed recall	63.0 (11.3)	77.2 (14.4)	76.5 (17.1)	103.7 (10.8)
HVLT-R				
Total recall	86.6 (13.4)	94.7 (14.1)	93.2 (13.0)	112.4 (13.6)
Delayed recall	72.1 (12.6)	86.9 (16.0)	85.0 (16.9)	108.0 (9.6)
COWAT	38.3 (13.1)	38.0 (13.0)	39.9 (10.1)	40.1 (10.6)
Animals	13.8 (4.3)	13.3 (4.2)	17.6 (4.1)	18.3 (4.7)
TMT-A	53.0 (15.6)	49.7 (15.9)	43.2 (18.3)	38.0 (10.9)
TMT-B	166.7 (68.9)	149.1 (76.9)	100.3 (31.8)	94.2 (32.8)
SDMT	32.0 (8.8)	32.9 (9.9)	41.3 (8.5)	42.3 (8.6)

BVMT-R = Brief Visuospatial Memory Test – Revised; HVLT-R = Hopkins Verbal Learning Test – Revised; COWAT = Controlled Oral Word Association Test; TMT = Trail Making Test; SDMT = Symbol Digit Modalities Test. Raw scores are presented for each measure, except BVMT-R and HVLT-R, which are age-corrected standard scores (M = 100, SD = 15). TMT scores are not reversed, as described in the Methods section in the calculation of practice effects scores. Test-rest correlations for each group can be obtained from the correpsonding author.

Int Psychogeriatr. Author manuscript; available in PMC 2013 October 03.

Table 6

Ratios of practice effects in participants reclassified after one week as MCI Stable, MCI-Intact, or Intact Stable.

PRACTICE EFFECTS	MCI STABLE	MCI-INTACT	INTACT STABLE
BVMT-R			
Total recall	1.14 (0.22)	1.37 (0.21)	1.31 (0.19)
Delayed recall	1.24 (0.21)	1.41 (0.31)	1.14 (0.14)
HVLT-R			
Total recall	1.10 (0.12)	1.21 (0.14)	1.13 (0.11)
Delayed recall	1.22 (0.25)	1.31 (0.24)	1.09 (0.12)
COWAT	1.01 (0.23)	1.03 (0.21)	1.06 (0.19)
Animals	0.99 (0.30)	1.05 (0.20)	1.07 (0.30)
TMT-A	1.05 (0.22)	1.05 (0.26)	1.06 (0.18)
TMT-B	1.09 (0.28)	1.03 (0.28)	1.07 (0.25)
SDMT	1.03 (0.12)	1.03 (0.10)	1.09 (0.16)

Practice effects were calculated as: one-week raw score/baseline raw score, except BVMT-R and HVLT-R, which are age-corrected standard scores (M=100, SD=15). BVMT-R = Brief Visuospatial Memory Test – Revised; HVLT-R = Hopkins Verbal Learning Test – Revised; COWAT = Controlled Oral Word Association Test; TMT = Trail Making Test; SDMT = Symbol Digit Modalities Test.*$p < 0.01$ after controlling for age differences.