HHS Public Access

Author manuscript

Clin Neuropsychol. Author manuscript; available in PMC 2023 August 01.

Published in final edited form as:

Clin Neuropsychol. 2022 August ; 36(6): 1304–1327. doi:10.1080/13854046.2020.1807058.

Validation of one-year reliable change in the RBANS for community-dwelling older adults with amnestic Mild Cognitive Impairment

Dustin B. Hammers[1,2], **Kayla R. Suhrie**[1], **Sariah M. Porter**[1], **Ava M. Dixon**[1], **Kevin Duff**[1,2]

[1]Center for Alzheimer's Care, Imaging, and Research, Department of Neurology, University of Utah

[2]Center on Aging, University of Utah

Abstract

OBJECTIVE: The current study sought to externally validate previously published standardized regression-based (SRB) equations for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) Indexes (Hammers, Suhrie, Porter, Dixon, & Duff, 2020) administered twice over a one-year period.

METHOD: Hammers and colleagues' SRB prediction equations were applied to two independent samples of community-dwelling older adults with amnestic Mild Cognitive Impairment (MCI), including those recruited from the community (n = 64) and those recruited from a memory disorders clinic (n = 58).

RESULTS: While Observed Baseline and Observed Follow-up performances were generally comparable for both MCI samples over one year, both samples possessed significantly lower Observed One-Year Follow-up scores than were predicted based on Hammers et al.'s development sample across many RBANS Indexes. Relatedly, both amnestic MCI samples possessed a greater percentage of participants either "declining" or failing to exhibit a long-term practice effect over one year relative to expectation across most Indexes. Further, the clinic-recruited amnestic MCI sample displayed worse baseline performances, smaller long-term practice effects, and greater proportions of individual participants exhibiting a decline across one year relative to the community amnestic MCI sample.

CONCLUSIONS: These findings validate Hammers et al.'s SRB prediction equations by (1) indicating their ability to identify clinically meaningful change across RBANS Indexes in

Address correspondences to Dustin B. Hammers, PhD, Center for Alzheimer's Care, Imaging and Research, Department of Neurology, University of Utah, 650 Komas Drive #106-A, Salt Lake City, UT 84108, USA. Tel: 801-585-3929. Fax: 801-581-2483. dustin.hammers@hsc.utah.edu.

Kayla Suhrie, Center for Alzheimer's Care, Imaging, and Research, Department of Neurology, University of Utah, 650 Komas Drive, Suite 106-A, Salt Lake City, UT 84108, USA, (801) 585-6546

Sariah Porter, Center for Alzheimer's Care, Imaging, and Research, Department of Neurology, University of Utah, 650 Komas Drive, Suite 106-A, Salt Lake City, UT 84108, USA, (801) 585-6546

Ava Dixon, Center for Alzheimer's Care, Imaging, and Research, Department of Neurology, University of Utah, 650 Komas Drive, Suite 106-A, Salt Lake City, UT 84108, USA, (801) 585-6546

Kevin Duff, Center for Alzheimer's Care, Imaging, and Research, Department of Neurology, University of Utah, 650 Komas Drive, Suite 106-A, Salt Lake City, UT 84108, USA, (801) 585-6546

Declaration of Interest Statement: No authors have any conflicts to declare.

independent samples, and (2) discriminating rates of cognitive change among cognitively nuanced samples.

Keywords

Cognition; Reliable Change; Assessment; Mild Cognitive Impairment; Memory

INTRODUCTION

Serial or longitudinal neuropsychological evaluation permits clinicians to monitor cognitive changes over time. However, care must be taken to ensure that any change observed between two testing sessions reflects clinically-relevant improvement or decline, as compared to statistical or methodological artifact (e.g., normal variation in performance, regression to the mean, testing variance, practice effect, etc.; Duff, 2012; Hammers, Duff, & Chelune, 2015; Lezak, Howieson, Bigler, & Tranel, 2012). To assist clinicians in determining the clinical meaningfulness of observed changes, statistical procedures known as *reliable change methods* have been developed (Chelune, 2003; Hinton-Bayre, 2010) and have gained wide acceptance over time (Attix et al., 2009; Crockford et al., 2018; Duff, 2014; Duff, Beglinger, Moser, Paulsen, et al., 2010; Gavett, Ashendorf, & Gurnani, 2015; Rinehardt et al., 2010; Sanchez-Benavides et al., 2016; Stein, Luppa, Brahler, Konig, & Riedel-Heller, 2010). Of the various approaches, McSweeny and colleagues' (McSweeny, Naugle, Chelune, & Luders, 1993) standardized regression-based (SRB) approach has been proposed as predicting cognitive change over time with the greatest accuracy (Duff, 2012).

Specifically, McSweeny's model utilizes multiple regression to predict individual performance at Time 2 (T_2) based on Time 1 (T_1) performance and various demographic characteristics (age, education, sex, etc.) and test conditions (e.g., retest interval). A discrepancy change score, or z score, can be calculated by subtracting the predicted T_2 performance (T_2') from the observed T_2 performance for a given individual, and normalizing the deviation using the standard error of the estimate of the regression (SE_{est}) via the following formula: $z = (T_2 - T_2')/ SE_{est}$. When attempting to interpret these calculations, positive z scores reflect an observed T_2 performance that was greater than predicted based on the SRB equations, whereas negative z scores reflect an observed T_2 performance that was lower than predicted. When using a 90% confidence interval of stability, which is standard for McSweeny's method (1993), a z of 1.645 equates to significance at an α value of $\alpha = .10$. As such, z scores above 1.645 commonly represent "improvement" (or benefit from practice beyond expectation), z scores below −1.645 commonly represent "decline" (or smaller practice effect relative to expectation), and z scores between +/− 1.645 commonly represent "stability" (or an anticipated level of benefit from practice). When considering the expected frequency of performance, 5% of participants would be expected to display "improvement", 90% would be expected to remain "stable," and 5% would be expected to display "decline" if the z scores were normally distributed. Considering conditions like Alzheimer's disease (AD) or Mild Cognitive Impairment (MCI) that are known to decline over time, the use of both baseline performance and relevant demographic and test-related characteristics has been suggested to

increase sensitivity towards identifying individuals that are at risk for a declining cognitive trajectory (Duff et al., 2017).

Duff and colleagues (Duff et al., 2004; Duff et al., 2005) were the first to calculate SRB prediction equations for the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 2012). They administered the RBANS twice over a one-year period to 445 cognitively intact community dwelling older adults recruited from primary care clinics, developing and validating SRB algorithms for RBANS Indexes of Immediate Memory, Visuospatial/Constructional, Language, Attention, Delayed Memory, and Total Scale score (Duff et al., 2004) and associated subtests (Duff et al., 2005). Because Duff's development and validation samples were dependent, Hammers and colleagues (Hammers, Suhrie, Porter, et al., 2020) attempted to externally validate these prediction algorithms using an independent sample of 129 robustly cognitively intact community-dwelling older adults. Results from this study suggest that the robustly intact sample possessed significantly better observed T_2 scores than were predicted based on Duff's development sample across most RBANS Indexes and many subtests. Since Duff's SRBs were not replicated in this robustly intact sample, Hammers et al. calculated updated prediction algorithms for the RBANS Index and subtest scores from those 129 robustly cognitively intact participants (see Hammers, Suhrie, Porter, et al., 2020).

These updated SRB prediction equations possess potential to provide more accurate assessment of reliable change in an individual patient because they were developed from a sample of cognitively intact older adults who remained stable over 12 months (Hammers, Suhrie, Porter, et al., 2020). The use of "clean" or "robust" normative samples is receiving heightened focus in recent years (Goodwill et al., 2019; Harrington et al., 2017), suggesting the importance of normative samples being "truly normal". However, as of yet Hammers et al.'s RBANS SRB equations have been unvalidated. Consequently, the primary aim of the current study was to evaluate the validity of these SRB prediction equations using two different samples of community-dwelling older adults with amnestic MCI, to allow for greater generalizability of Hammers et al.'s RBANS prediction equations to populations at risk for developing AD later in life. It was hypothesized that the application of these prediction equations to independent samples of older adults with amnestic MCI would result in a greater proportion of participants "declining" on these cognitive measures over one year than expected for each domain.

A secondary aim was to extend criterion validity by comparing the two amnestic MCI samples based on recruitment source (community versus clinic) on the amount of change observed over one year. Research has repeatedly shown that MCI participants recruited from memory clinics possess greater severity along the AD continuum relative to community-recruited MCI samples, including worse hippocampal volume declines (Whitwell et al., 2012), higher rates of apolipoprotein E4 allele findings (Brodaty et al., 2014), worse memory performance (Brodaty et al., 2014; Kirsebom et al., 2017), and greater risk of progression to dementia (Farias, Mungas, Reed, Harvey, & DeCarli, 2009; Roh et al., 2016). It was therefore hypothesized that greater frequencies of cognitive "decline" would be observed in the clinic-recruited amnestic MCI sample than the community-recruited amnestic MCI sample, as seen by greater discrepancies (z scores) between observed and

predicted T_2 performance for the clinic amnestic MCI sample. Together, validation of these SRB prediction equations could provide diagnostic and prognostic value, and inform treatment recommendations in patients with amnestic MCI based on research indicating that MCI patients who fail to benefit from practice on serial assessment have greater risk of AD-related pathology (Duff et al., 2018; Duff, Foster, & Hoffman, 2014; Galvin et al., 2005; Mormino et al., 2014), worse response to intervention (Duff, Beglinger, Moser, Schultz, & Paulsen, 2010), and worse outcomes (Duff et al., 2011; Hassenstab et al., 2015; Machulda et al., 2013).

METHOD

Participants

The current study was based on two different samples of participants (see Table 1). All participants were classified as having either single-domain or multi-domain amnestic MCI based on the protocol below. The first amnestic MCI sample, hereafter referred to as the "community amnestic MCI sample", included 64 community-dwelling older adults, recruited from senior centers and independent living facilities. The majority of this sample was recruited from 2010 to 2013 for a study of practice effects and amnestic MCI (see (Duff et al., 2017), with a minority recruited from 2016 to 2018 for a study of cognitive training and amnestic MCI (Duff et al., Under Review). Their mean age was 79.2 (SD = 7.9, range = 65 – 94) years old and they averaged 15.6 (SD = 2.6, range = 12 – 20) years of education. The majority of participants were female (76%) and Caucasian (100%), and premorbid intellect at baseline was at the upper limit of average according to the Wide Range Achievement Test –fourth edition (WRAT-4; Wilkinson & Robertson, 2006) Reading subtest (standard score: M = 109.8, SD = 10.8, range = 82 – 145).

The second amnestic MCI sample, hereafter referred to as the "clinic amnestic MCI sample", included 58 community-dwelling older adults, recruited from a cognitive disorder clinic from 2016 to 2018 for a study of cognitive training and amnestic MCI (Duff et al., Under Review). Their mean age was 74.1 (SD = 5.2, range = 66 – 88) years old and they possessed on average 17.0 (SD = 2.9, range = 12 – 20) years of education. Similar to the first sample, the majority of participants were Caucasian (98%), though 36% of the sample was female. Premorbid intellect at baseline was also average according to the WRAT-4 Reading subtest (standard score: M = 107.2, SD = 7.9, range = 91 – 123).

Classification of participants from both samples has been described previously (Duff, Dalley, Suhrie, & Hammers, 2019). Briefly, participants were classified as amnestic MCI based on criteria by Albert and colleagues (Albert et al., 2011) and Petersen (Petersen, 2004), which incorporated participant and knowledgeable informant report and a previously administered baseline cognitive evaluation. This previously administered evaluation included the Hopkins Verbal Learning Test – Revised (HVLT-R; Brandt & Benedict, 2001), the Brief Visuospatial Memory Test – Revised (BVMT-R; Benedict, 1997), Symbol Digit Modalities Test (SDMT; Smith, 1973), and the Trail Making Test Parts A and B (TMT-A and TMT-B; Reitan, 1992). Cognitive impairment for a domain was defined as a performance below the 7th percentile (SD = −1.5) relative to premorbid intellect, meaning that a difference of 22.5 *Standard Score* (SS) points between WRAT performance and a relevant task performance

was necessary for impairment. For example, if a participant's WRAT performance was SS = 100, then she/he would have been classified as being impaired for a task if her/his task performance was below SS = 77.5. Conversely, if a participant's WRAT performance was SS = 120, then she/he would have been classified as being impaired for a task if her/his task performance was below SS = 97.5. Inclusion criteria for both samples included being aged 65 years or older and functionally independent. Exclusion criteria for both samples included neurological conditions likely to affect cognition, dementia, major psychiatric condition, current severe depression, substance abuse, anti-convulsant or anti-psychotic medications, and residence in a skilled nursing or living facility. Finally, all participants in these two samples were classified as having amnestic MCI, therefore failure to exhibit memory impairment relative to premorbid intellect was an additional exclusion.

Hammers et al.'s (2020) SRB equation development sample – for which the RBANS Index and subtest SRB prediction equations were calculated – included 129 robustly cognitively intact community-dwelling older adults with a mean age of 75.6 (SD = 7.5) years and 15.4 (SD = 2.7) years of education. The sample of participants was 99% Caucasian and 83% female. Premorbid intellect at baseline was average according to the Wide Range Achievement Test – third or fourth editions (WRAT-3 and WRAT-4) Reading subtest (standard score: M = 107.4, SD = 7.6). These data are also summarized in Table 1.

Procedure

All procedures were approved by the local Institutional Review Board before the study commenced. All participants provided informed consent before completing any procedures. All participants were administered the RBANS (Randolph, 2012) at both baseline and approximately one year later. The RBANS is a neuropsychological test battery comprising 12 subtests that are used to calculate Index scores for domains of immediate memory, visuospatial/constructional, attention, language, delayed memory, and global neuropsychological functioning. Administration time for the RBANS is approximately 25–30 minutes. The Index scores utilize age-corrected normative comparisons from the test manual to generate standard scores (M = 100, SD = 15), with higher scores indicating better cognition. The subtest scores are represented as raw scores. To be consistent with the original Hammers et al. (2020) SRB equations, the same form of each test (Form A) was used across the two testing sessions.

Analyses

Demographic and Baseline Performance Analyses—Independent samples t tests were used to compare continuous demographic (e.g., age and education), testing (i.e., retest interval), and baseline cognitive performance variables between our community and clinic amnestic MCI samples, and one-sample t tests were used to compare continuous demographic, testing, and baseline performance variables for each amnestic MCI sample with Hammers et al.'s (2020) original sample. Additionally, two-way chi-square analyses were conducted between the two amnestic MCI samples to compare categorical demographic variables (e.g., sex, ethnicity), and one-sample chi-square analyses were conducted between each amnestic MCI sample and Hammers et al.'s original sample to compare categorical demographic variables.

Traditional Baseline vs. One-Year Follow-up Analyses—The Observed Baseline and One-Year Observed Follow-up scores for the RBANS Indexes were compared using pair-wise *t* tests for each of the amnestic MCI samples. These analyses were conducted to approximate a traditional evaluation of change over time (direct comparison of T_1 and T_2 scores) without controlling for long-term practice effects or participant variables.

SRB Group Analyses—Previously published SRB prediction equations for the RBANS Index scores were applied to the current RBANS Baseline and Follow-up performances of our community and clinic amnestic MCI samples. As has been described previously (Hammers, Suhrie, Porter, et al., 2020), the SRB prediction algorithms were calculated from a development sample using hierarchical multiple-regression analyses to maximize the prediction of RBANS performance. Specifically, the combination of demographic variables (e.g., age, education, sex, ethnicity), retest interval, and baseline RBANS Index score was used to predict the RBANS Index score at follow-up one year later. Age and education were represented as years old at baseline and number of years of formal education, respectively. Sex was coded as male = 0, and female = 1. Ethnicity was coded as non-White/non-Caucasian = 0, or White/Caucasian = 1. The retest interval was represented as days from T_1 to T_2. Index scores were age-corrected standard scores calculated from the RBANS manual, with a mean of 100 and a standard deviation of 15. Please see Table 2 for the specific SRB prediction algorithms for the RBANS Indexes (Hammers, Suhrie, Porter, et al., 2020) applied to the current samples.

Following the application of these SRB prediction equations to the current MCI samples' baseline RBANS Index scores and relevant demographic and testing characteristics, a *z* score was calculated for each participant, which reflects an individual normalized deviation of change. As indicated previously, the Observed One-Year Follow-up score was compared to the Predicted Follow-up score, normalized by the SE_{est} of the regression (McSweeny et al., 1993). While some discussion in the literature exists regarding the proper standard error estimate for use in reliable change methods (Hinton-Bayre, 2010), we have previously shown the equivalence of the two most-common estimates and provided support for use of the SE_{est} (Hammers & Duff, 2019). This *z* score was represented by the equation, $z = (T_2 - T_2')/ SE_{est}$. *Z* scores for each RBANS Index for both the community and clinic amnestic MCI samples were separately compared to expectation ($z = 0$) based on the normal distribution of *z* scores using a one-sample *t* test. Additionally, the resultant *z* scores were compared between the community and clinic amnestic MCI samples using independent samples *t* tests.

Individual Distribution Analyses—Finally, the resultant *z* scores were trichotomized into "decline", "stable", or "improve". As described above, if the *z* scores were normally distributed, then one would expect that 5% of participants would possess a *z* score < −1.645," 90% would possess a *z* score between +/− 1.645, and 5% would possess a *z* score > 1.645. Using this trichotomization, individual one-sample chi-square analyses were conducted for each RBANS Index score separately for each amnestic MCI sample to determine if the observed distribution of participants significantly deviated from the expected distribution based on the normal distribution of z scores. Additionally, two-way

chi-square analyses were conducted for the RBANS scores to determine if differences were observed between the community and clinic amnestic MCI samples in the distribution of participants who “declined”/remained “stable”/“improved”.

Measures of effect size were expressed throughout as Cohen’s *d* values for continuous data and *Phi* coefficients for categorical data. A two-tailed alpha level was set at .01 for all statistical analyses.

RESULTS

Demographic and Baseline Performance Analyses

Table 1 reflects demographic characteristics of participants from Hammers et al.’s (2020) development sample and the current community and clinic amnestic MCI validation samples. In the current samples, the community amnestic MCI sample was older, less educated, more predominantly female, and possessed shorter retest intervals on average than the clinic amnestic MCI sample (all *ps* < .001). Although differences were observed, these variables were all incorporated into the original Hammers et al. SRB equations, therefore their covariation in the subsequent recruitment group comparison analyses was not necessary. No differences were observed between samples for premorbid intellect ($p = .13$) or ethnicity ($p = .29$). When compared to Hammers et al.’s original development sample, as can be seen in Table 1, the community amnestic MCI sample was older and possessed a greater proportion of males than Hammers et al.’s development sample (both *ps* < .001). The clinic amnestic MCI sample was more educated, and possessed a longer retest interval and greater proportions of males than Hammers et al.’s development sample (all *p*s < .001). Expanded results (e.g., *t* values, effect sizes) are available upon request.

When considering Observed Baseline RBANS Index performances between the two amnestic MCI samples, differences were seen across nearly all Indexes. As can be seen in Table 1, Observed Baseline performances were lower for the clinic amnestic MCI sample relative to the community amnestic MCI sample for all RBANS Indexes (*p*s < .001) except for the Visuospatial/Constructional ($p = .56$) and Attention ($p = .76$) Indexes. Similarly, Observed Baseline performances for the clinic amnestic MCI sample were lower for all RBANS Indexes than Hammers et al.’s development sample (*p*s = .001), and also for the community amnestic MCI sample relative to the development sample for all RBANS Indexes (*p*s = .001) except Visuospatial/Constructional ($p = .01$; see Table 1). Expanded results are available upon request.

Traditional Baseline vs. One-Year Follow-up Analyses

Community amnestic MCI Sample—When examining change over time using a traditional method of comparing Observed Baseline and Observed One-Year Follow-up scores for the RBANS Indexes for the community amnestic MCI sample (see Table 3 for Means and *SD*s), significant differences were observed for the Indexes of Immediate Memory, $t(63) = -3.37$, $p = .001$, $d = -0.85$, and Delayed Memory, $t(63) = -4.28$, $p = .001$, $d = -1.08$. Specifically, the Observed One-Year Follow-up scores for both Indexes were higher than at Observed Baseline. No significant differences were observed for RBANS Indexes of

Visuospatial/Constructional, $t(63) = 1.89$, $p = .06$, $d = 0.48$, Language, $t(63) = 1.64$, $p = .11$, $d = 0.41$, Attention, $t(63) = -1.28$, $p = .21$, $d = -0.32$, or Total Scale score, $t(63) = -2.04$, $p = .05$, $d = -0.51$.

Clinic amnestic MCI Sample—For the clinic amnestic MCI sample (see Table 4), no significant differences were observed between Observed Baseline and Observed One-Year Follow-up for any of the RBANS Index scores, Immediate Memory, $t(57) = 0.61$, $p = .54$, $d = 0.16$, Visuospatial/Constructional, $t(57) = -0.11$, $p = .92$, $d = -0.03$, Language, $t(57) = 1.70$, $p = .10$, $d = 0.45$, Attention, $t(57) = 1.97$, $p = .05$, $d = 0.52$, Delayed Memory, $t(57) = 0.58$, $p = .56$, $d = 0.15$, or Total Scale score, $t(57) = 1.28$, $p = .21$, $d = 0.34$.

Traditional Baseline vs. One-Year Follow-up Summary—Observed Baseline scores were comparable to Observed One-Year Follow-up scores for the vast majority of RBANS Index scores. The only exceptions were seen in the community amnestic MCI sample, where the Observed One-Year Follow-up score was *higher* than the Observed Baseline score for the Indexes of Immediate Memory and Delayed Memory.

SRB Group Analyses

Community amnestic MCI Sample—SRB prediction equations for the RBANS Indexes from Hammers and colleagues (Hammers, Suhrie, Porter, et al., 2020) were applied to the current sample of community amnestic MCI participants. As seen in Table 3, when using one-sample *t* tests to compare *z* scores for each RBANS Index to an expected *z* score of zero based on the normal distribution of z scores, statistical significance was observed for 3 of 5 RBANS Indexes (Immediate Memory, $t(63) = -2.81$, $p = .007$, $d = -0.71$, Visuospatial/Constructional, $t(63) = -2.73$, $p = .008$, $d = -0.69$, and Language, $t(63) = -4.82$, $p = .001$, $d = -1.21$) and the Total Scale score, $t(63) = -3.07$, $p = .003$, $d = -0.77$. In each case, the resultant *z* score was significantly lower than zero. As a reminder, a negative *z* score indicates that the Observed Follow-up score was lower than the Predicted Follow-up score and suggested a reliable "decline" over one year, whereas a positive *z* score indicates that the Observed Follow-up score was higher than the Predicted Follow-up score suggested a reliable "improvement" over one year. Neither Indexes of Attention, $t(63) = -1.73$, $p = .09$, $d = -0.46$, nor Delayed Memory, $t(63) = -2.43$, $p = .02$, $d = -0.61$, were significant.

Clinic amnestic MCI Sample—For the clinic amnestic MCI participants (Table 4), when using one-sample *t* tests to compare *z* scores for each RBANS Index to an expected *z* score of zero, statistical significance was observed for all 5 RBANS Indexes (Immediate Memory, $t(57) = -10.42$, $p = .001$, $d = -2.76$, Visuospatial/Constructional, $t(57) = -3.00$, $p = .004$, $d = -0.79$, Language, $t(57) = -8.53$, $p = .001$, $d = -2.26$, Attention, $t(57) = -4.64$, $p = .001$, $d = -1.23$, Delayed Memory, $t(57) = -8.40$, $p = .001$, $d = -2.23$) and the Total Scale score, $t(57) = -8.89$, $p = .001$, $d = -2.36$. As with the community amnestic MCI sample, in each case the resultant *z* score was significantly lower than zero.

Community amnestic MCI Sample vs. Clinic amnestic MCI Sample—Further, when comparing *z* scores for each RBANS Index score between the community and clinic amnestic MCI samples (Tables 3 and 4) using independent samples *t* tests, significant

differences were observed for 3 of 5 Indexes (Immediate Memory, $t(120) = 5.72$, $p = .001$, $d = 1.04$, Language, $t(120) = 3.02$, $p = .003$, $d = 0.55$, Delayed Memory, $t(120) = 4.48$, $p = .001$, $d = 0.82$) and the Total Scale score, $t(120) = 3.64$, $p = .001$, $d = 0.66$. Specifically, z scores were significantly lower for the clinic amnestic MCI group than the community amnestic MCI group for all Indexes. No differences were observed between amnestic MCI samples for the RBANS Indexes of Visuospatial/Constructional, $t(120) = 0.01$, $p = .99$, $d = 0.01$, or Attention, $t(120) = 2.29$, $p = .02$, $d = 0.42$.

SRB Group Analyses Summary—After applying Hammers et al.'s (2020) SRB prediction equations to the community amnestic MCI sample, their z scores were significantly lower than zero for 3 of 5 RBANS Indexes (Immediate Memory, Visuospatial/Constructional, Language) and the Total Scale score. Z scores for the clinic amnestic MCI sample were lower than zero for all 5 RBANS Indexes and the Total Scale score. These negative z scores mean that the Observed One-Year Follow-up scores were lower than the Predicted One-Year Follow-up scores, suggesting a reliable "decline" over one year. When comparing between amnestic MCI samples, z scores were significantly lower for the clinic amnestic MCI group than the community amnestic MCI group for the Indexes of Immediate Memory, Language, Delayed Memory, and the Total Scale score.

Individual Distribution Analyses

When examining the distribution of community amnestic MCI and clinic amnestic MCI participants that either "declined", remained "stable", or "improved" relative to predictions over the one-year interval between Baseline and One-Year Follow-up administrations of the RBANS, the majority of participants remained "stable" (86% of community MCI participants and 75% of clinic MCI participants) for both the RBANS Indexes (see Table 5).

Community amnestic MCI Sample—For the community amnestic MCI sample, using one-sample chi square analyses, statistically different distributions were observed relative to what was anticipated based on the normal curve distribution of z scores (i.e., 5% "decline", 90% remained "stable", 5% "improve") for 1 of 5 RBANS Indexes (Delayed Memory, χ^2 (2) = 15.22, $p = .001$, $Phi = 0.49$). Specifically, greater proportions of individuals "declined" than expected based on normal distributions for this Index (16% "declined" on Delayed Memory; see Table 5). No significant differences in distributions from anticipation were observed for the Indexes of Immediate Memory, χ^2 (2) = 4.75, $p = .09$, $Phi = 0.27$, Visuospatial/Constructional, χ^2 (2) = 4.75, $p = .09$, $Phi = 0.27$, Language, χ^2 (2) = 7.88, $p = .02$, $Phi = 0.35$, Attention, χ^2 (2) = 0.50, $p = .78$, $Phi = 0.09$, or Total Scale score, χ^2 (2) = 2.16, $p = .34$, $Phi = 0.18$.

Clinic amnestic MCI Sample—Relatedly, for the clinic amnestic MCI sample, one-sample chi square analyses revealed that statistically different distributions were observed relative to anticipation based on the normal curve distribution of z scores for 3 of 5 RBANS Indexes (Immediate Memory, χ^2 (2) = 211.79, $p = .001$, $Phi = 1.91$, Language, χ^2 (2) = 20.61, $p = .001$, $Phi = 0.60$, Delayed Memory, χ^2 (2) = 210.97, $p = .001$, $Phi = 1.91$) and Total Scale score, χ^2 (2) = 20.61, $p = .001$, $Phi = 0.60$. Specifically, greater proportions of individuals "declined" than expected based on normal distributions for each Index (17% of

participants “declined” on both the Language and Total Scale score, and 47% “declined” on both the Immediate Memory and Delayed Memory). No significant differences in distributions from anticipation were observed for the Indexes of Visuospatial/Constructional, χ^2 (2) = 2.77, p = .25, *Phi* = 0.22, and Attention, χ^2 (2) = 2.77, p = .25, *Phi* = 0.22.

Community amnestic MCI Sample vs. Clinic amnestic MCI Sample—When comparing distributions of “decline”/“stable”/“improve” between the community amnestic MCI and clinic amnestic MCI samples over one year, two-way chi-square analyses revealed significant differences for 2 of 5 RBANS Indexes (Immediate Memory, χ^2 (2) = 76.59, p = .001, *Phi* = 0.79, and Delayed Memory, χ^2 (2) = 42.28, p = .001, *Phi* = 0.59) and the Total Scale score, χ^2 (2) = 38.92, p = .001, *Phi* = 0.56. Specifically, a greater percentage of participants “declined” on the Immediate Memory and Delayed Memory Indexes, and Total Scaled score, for the clinic amnestic MCI sample compared to the community amnestic MCI sample (47% versus 11%, respectively, for Immediate Memory, 47% versus 16%, respectively, for Delayed Memory, and 17% versus 3%, respectively, for Total Scale score). No differences were observed between the community and clinic amnestic MCI samples for Indexes of Visuospatial/Constructional, χ^2 (2) = 1.56, p = .46, *Phi* = 0.11, Language, χ^2 (2) = 2.86, p = .24, *Phi* = 0.15, or Attention, χ^2 (2) = 2.32, p = .31, *Phi* = 0.14.

Individual Distribution Analyses Summary—Although most participants remained within the “stable” range on repeat testing, greater proportions of individuals “declined” over one year for the community amnestic MCI sample relative to expectation for RBANS Delayed Memory Index. Greater proportions of individuals “declined” over one year for the clinic amnestic MCI sample relative to expectation for Indexes of Immediate Memory, Language, Delayed Memory, and Total Scale score. When comparing the two amnestic MCI samples, a greater percentage of participants “declined” on the Immediate Memory and Delayed Memory Indexes, and the Total Scale score, for the clinic amnestic MCI sample compared to the community amnestic MCI sample.

DISCUSSION

The current study sought to examine the validity of previously published SRB predicted difference equations for the RBANS Indexes (Hammers, Suhrie, Porter, et al., 2020) using two independent samples of older adults with amnestic MCI who were living independently in the community and assessed twice over a one-year period. These algorithms were previously calculated from 129 robustly cognitively intact older adults and possess the potential to improve the accuracy of one-year serial RBANS prediction, but they have yet to be validated. Additionally, the current study extended previous research by comparing community-recruited versus memory-clinic-recruited amnestic MCI samples on the amount of change observed over one year using reliable change methods, which to our knowledge is the first study to do so.

For our current validation samples of both community MCI and clinic amnestic MCI participants, no change in performance was observed over one year when using the traditional method of comparing observed test scores at baseline and one-year across most RBANS Indexes administered. The exceptions were the Immediate and Delayed Memory

Indexes for the community amnestic MCI sample, where participants' performance was on average better at one-year follow-up than at baseline. While potentially unexpected, the literature suggests a wide discrepancy of findings when examining rates of annual conversion from MCI to AD, ranging from 4–40% of patients declining to AD per year (Boyle, Wilson, Aggarwal, Tang, & Bennett, 2006; Gauthier et al., 2006; Panza et al., 2007; Panza et al., 2005). This suggests that some studies – like the current study – have observed stability over one year in MCI samples, whereas others have observed a moderate degree of decline.

In contrast, when applying Hammers et al.'s (2020) SRB prediction equations to baseline performance on the RBANS, the Observed One-Year Follow-up scores for both our samples of amnestic MCI participants were consistently *below expectation* compared to predictions, such that participants tended to score worse than predicted. First, *z* scores for the community amnestic MCI sample were significantly lower than anticipated for Indexes of Immediate Memory, Visuospatial/Constructional, Language, and the Total Scale score (Cohen's *d*s = | 0.69 – 1.21 |). For the clinic amnestic MCI sample, *z* scores were significantly lower than anticipated for all five Indexes and the Total Scale score (Cohen's *d*s = | 0.79 – 2.76 |). As indicated above, a *z* score being lower than the expected value of 0 (negative *z* score) indicates that the Observed Follow-up score was lower than the Predicted Follow-up score. These findings are consistent with the literature suggesting that the use of SRB prediction equations can be more sensitive to change over time than the traditional baseline versus follow-up assessment method (Duff, Suhrie, Dalley, Anderson, & Hoffman, 2019).

Second, when examining the distributions of participants that displayed "improvement" (or greater benefit from practice than expected), "stability" (or expected long-term practice effects), or "declines" (or smaller long-term practice effects) relative to predictions, greater proportions of individuals performed worse than expected based on normal distributions in each amnestic MCI sample for select Indexes. For example, for the community amnestic MCI sample, 16% of participants "declined" on the Delayed Memory Index. For the clinic amnestic MCI sample, 17% of participants "declined" on both the Language Index and Total Scale score, and 47% "declined" on both the Immediate Memory and Delayed Memory Indexes. On no Index did a significant number of participants "improve" beyond expectation relative to predictions (e.g., > 5–10% of the sample "improving" beyond predictions). Because predicted scores are generally expected to improve from baseline to follow-up due to the benefit from repeated exposure, the current collective performances suggest that our two amnestic MCI samples generally reflected a reliable "decline" relative to expectation for delayed memory and other cognitive domains, and they experienced lesser long-term practice effects than expected on the RBANS over one year. Overall, potential reasons that the one-year RBANS scores in the current samples are consistently lower than predicted include the following: (1) the equations are not representative or appropriate for the samples, or (2) the equations are appropriate and also permitting observation of the subtle decline over one year that is not observed through "traditional" methods. When examining Table 5 specifically, the level of variance (or the % of those Improving or Declining) does not appear to be occurring at random, but instead the Decliners (10% across tasks for community amnestic MCI, 24% for clinic amnestic MCI) greatly outweigh the Improvers (4% across tasks for community amnestic MCI, 1% for clinic amnestic MCI). This is consistent with

expectations for cognitively compromised samples, and suggests that our SRB equations are permitting observation of subtle decline. Taken together, the current results appear to add external support to the validity of Hammers et al.'s SRB equations for the RBANS Indexes to predict cognitive performance one year after baseline assessment.

Our reduced long-term practice effect finding is consistent with several studies reporting an absence or a reduction of practice effects in MCI across a number of cognitive measures and retest intervals (Britt et al., 2011; Calamia, Markon, & Tranel, 2012; Cooper, Lacritz, Weiner, Rosenberg, & Cullum, 2004; Darby, Maruff, Collie, & McStephen, 2002; Schrijnemaekers, de Jager, Hogervorst, & Budge, 2006). However, worse outcomes were not observed for all RBANS Indexes for these amnestic MCI samples based on expectation in our study, and overall 86% (community amnestic MCI) and 75% (clinic amnestic MCI) of participants still displayed the expected level of long-term practice effect relative to Hammers et al.'s prediction equations (2020). Therefore, long-term practice effects were not non-existent in these amnestic MCI samples. Several authors have previously provided supportive evidence on the presence of practice effects in MCI participants (Duff et al., 2007; Mathews et al., 2014; Yan & Dick, 2006), and it has been proposed that while declarative memory is impacted early in the course of MCI, procedural memory is expected to stay stable in most individuals with amnestic MCI until later in the condition (Duff et al., 2008; Yan & Dick, 2006). Consequently, we would propose that long-term practice effects appear to be impacted by both declarative and procedural memory, such that even patients further along the amnestic MCI continuum (i.e., those with severe memory impairment at baseline) still tend to benefit from practice to a certain extent. Also, as our study focused on samples with amnestic MCI and less on other non-memory domains, it is possible that the continued evidence of long-term practice effects may be influenced by other cognitive domains like executive functioning. Specifically, executive functioning has been implicated in influencing both short- and long-term practice effects by reducing the novelty-effect for tests (Suchy, Kraybill, & Franchow, 2011; Thorgusen, Suchy, Chelune, & Baucom, 2016), which may have contributed to our findings in these amnestic MCI samples.

Additionally, the results of analyses from our secondary aim appear to extend the validation of Hammers et al.'s (2020) SRB prediction equations. Specifically, we included two sets of amnestic MCI samples – those recruited from the community versus those recruited from a memory disorders clinic – to identify whether differential rates of cognitive change or long-term practice effect were observed across different recruitment samples of amnestic MCI participants. Our results indicated that not only did the clinic amnestic MCI sample perform worse at baseline relative to the community amnestic MCI sample for several of the RBANS Indexes (Immediate Memory, Language, Delayed Memory, and Total Scale score), they also tended to exhibit smaller long-term practice effects over one year – as measured by lower z scores for the clinic amnestic MCI sample than the community amnestic MCI sample for several RBANS Indexes (Immediate Memory, Language, Delayed Memory, and Total Scale score), and also by greater proportions of participants in the clinic amnestic MCI sample displaying "declines" from prediction over one year than the community amnestic MCI sample across many RBANS Indexes (Immediate Memory, Delayed Memory, and Total Scale score). Our baseline results are consistent with repeated observations that source of recruitment has been shown to influence findings using MCI participants. Specifically,

recruiting patients with MCI from memory clinics has resulted in worse performance on memory tasks (Brodaty et al., 2014; Kirsebom et al., 2017), more severe hippocampal atrophy (Whitwell et al., 2012), and greater risk of MCI-to-dementia progression (Farias et al., 2009; Roh et al., 2016). As a result, it appears that studies using MCI participants should be intentional about their choices regarding recruitment source, as patients may tend to present more severely on the AD continuum if recruited from a clinic.

Regarding the differential long-term practice effects observed between the two amnestic MCI samples, these findings are consistent with research suggesting lower practice effects tend to be observed in more severely cognitively compromised samples (Calamia et al., 2012; Cooper et al., 2004; Gavett et al., 2016). Specifically, we observed that z scores were consistently lower for the clinic amnestic MCI sample than the community amnestic MCI sample, and that Observed Baseline performances were also lower for the clinic amnestic MCI sample across RBANS Indexes. This suggests that despite incorporating Baseline performance in the SRB equations, differences in baseline performance were likely an important contributor to the observed long-term practice effect differences between our two amnestic MCI samples. Rapport (Rapport et al., 1997; Rapport, Brines, Axelrod, & Theisen, 1997) has previously used the phrase “the rich get richer” to describe this effect. When applied to our samples, our community amnestic MCI sample possessed stronger baseline performances on several domains of cognition, which appears to have left them poised to benefit from practice to a greater extent than the clinic amnestic MCI sample, leading to the relatively higher *z* scores for the community amnestic MCI sample. Additionally, the properties of most regression-based prediction formulas are such that greater accuracy of prediction occurs for individuals falling closer to the mid-range of the performance scale, versus falling at the extremes (Tabachnick & Fidell, 1996). As will be observed in an example below, this likely led to expected improvements at T_2 being higher for lower baseline performances (e.g., SS = 80) than for average baseline performances (e.g., SS = 100) in our sample. As a result of Hammers et al.’s SRBs being more susceptible to regression to the mean effects at the extremes, they likely over-estimated T_2 predictions for low baseline scores or under-estimated T_2 predictions for very high baseline scores in our sample. Subsequently, this suggests that future work examining practice effects in cognitive data should continue to consider the importance of baseline performance as an underlying mechanism. Overall, as a result of these recruitment sample comparisons, the current study supports that the Hammers et al. (2020) prediction equations are generally able to discriminate rates of cognitive change among cognitively nuanced samples.

An example may be helpful to highlight two potential patterns that can be observed when applying these Hammers et al. (2020) prediction equations to individual patients. Please see Table 6 for Observed Baseline, Observed One-Year Follow-up, and Predicted One-Year Follow-up scores for a 70-year-old female participant with 12 years of education, who was administered the RBANS twice over one year. The first pattern is that of a patient who appears to improve on an Index at follow-up by a large amount relative to baseline (i.e., traditional method of examining change), but the SRB prediction equations indicate that the change is not beyond expectation. Specifically, on the Language Index of the RBANS, our patient improves by 15 points from T_1 to T_2, which would appear to be a large change using the traditional method of change. However, using the SRB method, her *z* score for this

Index is only 1.17 due to the expected long-term practice effect benefit at T_2 (the Predicted Follow-up score was expected to improve by 3.44 points over Baseline performance). This z score is lower than the z ±1.645 threshold, consequently this performance does not reflect a statistically significant improvement despite the large discrepancy between observed baseline and follow-up scores. A second pattern can also emerge, in that a patient appears to decline only a small amount (i.e., traditional method), but when factoring in expected improvement, the performance at T_2 reflects a significant decline (via the SRB method). For example, our patient displayed a decline of 7 points in her Delayed Memory Index, which may be dismissed as insufficient change in the traditional method. However, using the SRB method, this amount of change is reflective of a z score on this Index of −1.70. This z score is higher than the z ±1.645 threshold, indicating that this performance does reflect a statistically significant decline. Within the SRB method, such a large negative z score reflects the expectation that Hammers et al.'s development sample benefited from practice at T_2 on the Delayed Memory Index. As a result, instead of our patient improving by 11.51 points (Predicted Follow-up relative to Baseline), she declined by 7 points, which led to the large discrepancy between Observed and Predicted Follow-up (18.51 points) and the subsequent large negative z score. This same pattern is observed on the Immediate Memory Index, where our patient declines by 2 points at T_2, her Predicted Follow-up score was expected to improve by 16.37 points relative to baseline. Her actual discrepancy between Observed and Predicted Follow-up on this Index is therefore large (18.37 points) and reflective of a z score of −1.69. As both of these latter z score values are greater than the z ±1.645 threshold, our patient displays statistically significant "declines" in her Immediate and Delayed Memory abilities over one year.

This case example points out an important distinction between statistical significance and clinical relevance when incorporating reliable change methods into clinical or research decisions. The RBANS Index z score values that reached statistical significance (Tables 3 and 4) ranged between −0.37 and −1.60 therefore these values would not have been significant (i.e., exceeding z = −1.645) on the individual level. However, by focusing on the proportion of individuals per sample possessing clinically meaningful "decline" in Observed relative to Predicted Follow-up scores, these latter results highlight the greater percentage of amnestic MCI patients "declining" or failing to observe long-term practice effects at one-year follow-up on the RBANS. Such findings lend support to the ability of Hammers and colleagues' SRB prediction equations (2020) of identifying clinically meaningful change in individual patients or participants.

The current study is not without limitations. First, these results only inform us about change in participants with amnestic MCI, and do not speak to the Hammers et al. (2020) SRB algorithms ability to predict change in patients with other disease states ranging from AD and Frontotemporal dementia to those being cognitively intact. Further consideration of validity and generalizability of these prediction equations should be withheld until future studies can examine patients with these conditions. Additionally, the lack of clinical characterization or identification of probable etiology of our amnestic MCI samples may be a limitation to possible interpretation and generalization of results. Second, some research has suggested that MCI is not as accurately captured by the RBANS compared to a more sensitive neuropsychological battery (Duff, Hobson, Beglinger, & O'Bryant, 2010),

therefore these findings may not generalize to other memory tests when working with amnestic MCI participants. Similarly, these findings were observed for retest intervals of approximately one year, and until further studies are conducted it cannot be determined if they generalize to other retest intervals (Calamia et al., 2012). Third, the original studies that generated both the Duff (Duff et al., 2004) and Hammers et al. (2020) SRBs used RBANS Form A for all evaluations, consequently Duff's and Hammers et al.'s SRBs are based on the same forms being used. Since we were attempting to validate these Hammers et al.'s SRBs, we were therefore limited by the original studies and administered RBANS Form A for all evaluations as well. Future studies should consider practice effects in alternate forms of the RBANS as a new area of investigation, particularly examining differences in cognitively normal and clinical stable individuals, as this more closely approximates clinical practice.

Fourth, it is unclear if the current results would be similarly observed in a study incorporating more heterogeneous participants in regards to premorbid functioning, education, ethnicity, sex, or baseline performance. For example, as both of our samples were classified as amnestic MCI, our baseline RBANS Index scores for Immediate Memory and Delayed Memory Indexes tended towards lower on the spectrum of performance. As a result, these performances likely fell below the mid-range of performance for Hammers et al.'s (2020) prediction equations, leading to increased susceptibility of regression to the mean and subsequent over-prediction of T_2 performance (Tabachnick & Fidell, 1996). This subsequent over-prediction, particularly for the Immediate Memory Index, may result in the appearance of "declines" relative to prediction for an individual who performed generally comparable between T_1 and T_2 (see Table 6). As such, diagnosticians should be mindful of the impact of this potential over-prediction if applying these prediction equations – especially for the Immediate Memory Index – for diagnostic or classification purposes. Additionally, both of our samples possessed elevated levels of premorbid intellect and education. While these levels were generally consistent with Hammers et al.'s development sample, they differed from the general population. As both premorbid intellect and education have been used as proxy measures for cognitive reserve (Jefferson et al., 2011), these factors have been suggested as protective factors in the development of dementia (Stern, 2006), influences on the onset of AD symptoms (Roe, Xiong, Grant, Miller, & Morris, 2008), and moderators in the development of AD pathology (Rentz et al., 2010; Rodrigue et al., 2012; Roe, Mintun, et al., 2008). Further, both of our samples were predominantly Caucasian, and the community sample was predominantly female. While future research should consider such predictions in samples that are not primarily well-educated Caucasian females, the current study's proportion of highly educated females in the community sample (along with Hammers et al.'s original development sample) appear to reflect long-standing trends in research participation. Specifically, it has been observed that women tend to volunteer more than men across all age ranges (United States Bureau of the Census, Statistics, National, & Service, 2015), reaching a difference of upwards of 30% (U.S. Bureau of Labor Statistics, 2016), and that individuals with higher education and Caucasians consistently volunteer at greater levels (United States Bureau of the Census et al., 2015). Together, future studies examining these SRBs across a wide range of RBANS performance, premorbid intellect, education levels, sex, and ethnicity are therefore encouraged.

Despite these limitations, the current results appear to externally validate the SRB prediction equations developed by Hammers and colleagues (2020), which can be used to more accurately quantify clinically meaningful change over one year when using the RBANS. Given the multitude of research suggesting the importance of practice effects on treatment response and benefit in MCI samples (Duff et al., 2018; Duff, Hobson, et al., 2010; Duff et al., 2011; Galvin et al., 2005; Hassenstab et al., 2015; Machulda et al., 2013; Mormino et al., 2014), these current results also support the potential of these SRB prediction algorithms to provide diagnostic and prognostic value, and inform treatment recommendations.

Acknowledgements:

The project described was supported by research grants from the National Institutes on Aging: K23 AG028417, 5R01 AG045163, and 5R01 AG055428. 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

Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, … Phelps CH (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia: The Journal of the Alzheimer's Association, 7(3), 270–279. doi:10.1016/j.jalz.2011.03.008

Attix DK, Story TJ, Chelune GJ, Ball JD, Stutts ML, Hart RP, & Barth JT (2009). The prediction of change: normative neuropsychological trajectories. Clin Neuropsychol, 23(1), 21–38. doi:10.1080/13854040801945078 [PubMed: 18720272]

Benedict R (1997). Brief visuospatial memory test - Revised: professional manual. Lutz, FL: Psychological Assessment Resources, Inc.

Boyle PA, Wilson RS, Aggarwal NT, Tang Y, & Bennett DA (2006). Mild cognitive impairment: risk of Alzheimer disease and rate of cognitive decline. Neurology, 67(3), 441–445. doi:10.1212/01.wnl.0000228244.10416.20 [PubMed: 16894105]

Brandt J, & Benedict R (2001). Hopkins Verbal Learning Test - Revised. Odessa, FL: PAR.

Britt WG 3rd, Hansen AM, Bhaskerrao S, Larsen JP, Petersen F, Dickson A, … Kirsch WM (2011). Mild cognitive impairment: prodromal Alzheimer's disease or something else? J Alzheimers Dis, 27(3), 543–551. doi:10.3233/JAD-2011-110740 [PubMed: 21841247]

Brodaty H, Mothakunnel A, de Vel-Palumbo M, Ames D, Ellis KA, Reppermund S, … Sachdev PS (2014). Influence of population versus convenience sampling on sample characteristics in studies of cognitive aging. Ann Epidemiol, 24(1), 63–71. doi:10.1016/j.annepidem.2013.10.005 [PubMed: 24211070]

Calamia M, Markon K, & Tranel D (2012). Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. Clin Neuropsychol, 26(4), 543–570. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22540222 [PubMed: 22540222]

Chelune G (2003). Assessing reliable neuropsychological change. In R F (Ed.), Prediction in forensic and neuropsychology: New approaches to psychometrically sound assessment. Nahwah, NJ: Erlbaum.

Cooper DB, Lacritz LH, Weiner MF, Rosenberg RN, & Cullum CM (2004). Category fluency in mild cognitive impairment: reduced effect of practice in test-retest conditions. Alzheimer Dis Assoc Disord, 18(3), 120–122. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15494616 [PubMed: 15494616]

Crockford C, Newton J, Lonergan K, Madden C, Mays I, O'Sullivan M, … Abrahams S (2018). Measuring reliable change in cognition using the Edinburgh Cognitive and Behavioural ALS Screen (ECAS). Amyotroph Lateral Scler Frontotemporal Degener, 19(1–2), 65–73. doi:10.1080/21678421.2017.1407794 [PubMed: 29214872]

Darby D, Maruff P, Collie A, & McStephen M (2002). Mild cognitive impairment can be detected by multiple assessments in a single day. Neurology, 59(7), 1042–1046. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12370459 [PubMed: 12370459]

Duff et al. (Under Review). Brain plasticity-based training in amnestic Mild Cognitive Impairment: Treatment response and the impact of practice effects. Journal of Geriatric Psychiatry & Neurology.

Duff K (2012). Evidence-based indicators of neuropsychological change in the individual patient: relevant concepts and methods. Arch Clin Neuropsychol, 27(3), 248–261. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22382384 [PubMed: 22382384]

Duff K (2014). One-week practice effects in older adults: tools for assessing cognitive change. Clin Neuropsychol, 28(5), 714–725. doi:10.1080/13854046.2014.920923 [PubMed: 24882553]

Duff K, Anderson JS, Mallik AK, Suhrie KR, Atkinson TJ, Dalley BCA, … Hoffman JM (2018). Short-term repeat cognitive testing and its relationship to hippocampal volumes in older adults. J Clin Neurosci, 57, 121–125. doi:10.1016/j.jocn.2018.08.015 [PubMed: 30143414]

Duff K, Atkinson TJ, Suhrie KR, Dalley BC, Schaefer SY, & Hammers DB (2017). Short-term practice effects in mild cognitive impairment: Evaluating different methods of change. J Clin Exp Neuropsychol, 39(4), 396–407. doi:10.1080/13803395.2016.1230596 [PubMed: 27646966]

Duff K, Beglinger LJ, Moser DJ, Paulsen JS, Schultz SK, & Arndt S (2010). Predicting cognitive change in older adults: the relative contribution of practice effects. Arch Clin Neuropsychol, 25(2), 81–88. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20064816 [PubMed: 20064816]

Duff K, Beglinger LJ, Moser DJ, Schultz SK, & Paulsen JS (2010). Practice effects and outcome of cognitive training: preliminary evidence from a memory training course. Am J Geriatr Psychiatry, 18(1), 91. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20104658 [PubMed: 20104658]

Duff K, Beglinger LJ, Schultz SK, Moser DJ, McCaffrey RJ, Haase RF, … Huntington's Study, G. (2007). Practice effects in the prediction of long-term cognitive outcome in three patient samples: a novel prognostic index. Arch Clin Neuropsychol, 22(1), 15–24. doi:10.1016/j.acn.2006.08.013 [PubMed: 17142007]

Duff K, Beglinger LJ, Van Der Heiden S, Moser DJ, Arndt S, Schultz SK, & Paulsen JS (2008). Short-term practice effects in amnestic mild cognitive impairment: implications for diagnosis and treatment. Int Psychogeriatr, 20(5), 986–999. doi:10.1017/S1041610208007254 [PubMed: 18405398]

Duff K, Dalley BCA, Suhrie KR, & Hammers DB (2019). Predicting Premorbid Scores on the Repeatable Battery for the Assessment of Neuropsychological Status and their Validation in an Elderly Sample. Arch Clin Neuropsychol, 34(3), 395–402. doi:10.1093/arclin/acy050 [PubMed: 29878036]

Duff K, Foster NL, & Hoffman JM (2014). Practice effects and amyloid deposition: preliminary data on a method for enriching samples in clinical trials. Alzheimer Dis Assoc Disord, 28(3), 247–252. doi:10.1097/WAD.0000000000000021 [PubMed: 24614265]

Duff K, Hobson VL, Beglinger LJ, & O'Bryant SE (2010). Diagnostic accuracy of the RBANS in mild cognitive impairment: limitations on assessing milder impairments. Arch Clin Neuropsychol, 25(5), 429–441. doi:10.1093/arclin/acq045 [PubMed: 20570820]

Duff K, Lyketsos CG, Beglinger LJ, Chelune G, Moser DJ, Arndt S, … McCaffrey RJ (2011). Practice effects predict cognitive outcome in amnestic mild cognitive impairment. Am J Geriatr Psychiatry, 19(11), 932–939. doi:10.1097/JGP.0b013e318209dd3a [PubMed: 22024617]

Duff K, Schoenberg MR, Patton D, Mold J, Scott JG, & Adams RL (2004). Predicting change with the RBANS in a community dwelling elderly sample. J Int Neuropsychol Soc, 10(6), 828–834. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/15637773 [PubMed: 15637773]

Duff K, Schoenberg MR, Patton D, Paulsen JS, Bayless JD, Mold J, … Adams RL (2005). Regression-based formulas for predicting change in RBANS subtests with older adults. Arch Clin Neuropsychol, 20(3), 281–290. doi:10.1016/j.acn.2004.07.007 [PubMed: 15797165]

Duff K, Suhrie KR, Dalley BCA, Anderson JS, & Hoffman JM (2019). External validation of change formulae in neuropsychology with neuroimaging biomarkers: A methodological recommendation and preliminary clinical data. Clin Neuropsychol, 33(3), 478–489. doi:10.1080/13854046.2018.1484518 [PubMed: 29884099]

Farias ST, Mungas D, Reed BR, Harvey D, & DeCarli C (2009). Progression of mild cognitive impairment to dementia in clinic- vs community-based cohorts. Arch Neurol, 66(9), 1151–1157. doi:10.1001/archneurol.2009.106 [PubMed: 19752306]

Galvin JE, Powlishta KK, Wilkins K, McKeel DW Jr., Xiong C, Grant E, … Morris JC (2005). Predictors of preclinical Alzheimer disease and dementia: a clinicopathologic study. Arch Neurol, 62(5), 758–765. doi:10.1001/archneur.62.5.758 [PubMed: 15883263]

Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, … International Psychogeriatric Association Expert Conference on mild cognitive, i. (2006). Mild cognitive impairment. Lancet, 367(9518), 1262–1270. doi:10.1016/S0140-6736(06)68542-5 [PubMed: 16631882]

Gavett BE, Ashendorf L, & Gurnani AS (2015). Reliable Change on Neuropsychological Tests in the Uniform Data Set. J Int Neuropsychol Soc, 21(7), 558–567. doi:10.1017/S1355617715000582 [PubMed: 26234918]

Gavett BE, Gurnani AS, Saurman JL, Chapman KR, Steinberg EG, Martin B, … Stern RA (2016). Practice Effects on Story Memory and List Learning Tests in the Neuropsychological Assessment of Older Adults. PLoS One, 11(10), e0164492. doi:10.1371/journal.pone.0164492 [PubMed: 27711147]

Goodwill AM, Campbell S, Henderson VW, Gorelik A, Dennerstein L, McClung M, & Szoeke C (2019). Robust norms for neuropsychological tests of verbal episodic memory in Australian women. Neuropsychology, 33(4), 581–595. doi:10.1037/neu0000522 [PubMed: 30829514]

Hammers D, Duff K, & Chelune G (2015). Assessing change of cognitive trajectories over time in later life. In Pachana NA & Laidlaw K (Eds.), Oxford Handbook of Clinical Geropsychology. Oxford: Oxford University Press.

Hammers DB, & Duff K (2019). Application of Different Standard Error Estimates in Reliable Change Methods. Arch Clin Neuropsychol. doi:10.1093/arclin/acz054

Hammers DB, Suhrie KR, Dixon A, Porter S, & Duff K (2020). Reliable change in cognition over 1 week in community-dwelling older adults: a validation and extension study. Arch Clin Neuropsychol. doi:10.1093/arclin/acz076

Hammers DB, Suhrie KR, Porter SM, Dixon AM, & Duff K (2020). Generalizability of reliable change equations for the RBANS over one year in community-dwelling older adults. J Clin Exp Neuropsychol, 1–12. doi:10.1080/13803395.2020.1740654

Harrington KD, Lim YY, Ames D, Hassenstab J, Rainey-Smith S, Robertson J, … Group, A. R. (2017). Using Robust Normative Data to Investigate the Neuropsychology of Cognitive Aging. Arch Clin Neuropsychol, 32(2), 142–154. doi:10.1093/arclin/acw106 [PubMed: 27932344]

Hassenstab J, Ruvolo D, Jasielec M, Xiong C, Grant E, & Morris JC (2015). Absence of practice effects in preclinical Alzheimer's disease. Neuropsychology, 29(6), 940–948. doi:10.1037/neu0000208 [PubMed: 26011114]

Hinton-Bayre AD (2010). Deriving reliable change statistics from test-retest normative data: comparison of models and mathematical expressions. Arch Clin Neuropsychol, 25(3), 244–256. doi:10.1093/arclin/acq008 [PubMed: 20197293]

Jefferson AL, Gibbons LE, Rentz DM, Carvalho JO, Manly J, Bennett DA, & Jones RN (2011). A life course model of cognitive activities, socioeconomic status, education, reading ability, and cognition. J Am Geriatr Soc, 59(8), 1403–1411. doi:10.1111/j.1532-5415.2011.03499.x [PubMed: 21797830]

Kirsebom BE, Espenes R, Waterloo K, Hessen E, Johnsen SH, Brathen G, … Fladby T (2017). Screening for Alzheimer's Disease: Cognitive Impairment in Self-Referred and Memory Clinic-Referred Patients. J Alzheimers Dis, 60(4), 1621–1631. doi:10.3233/JAD-170385 [PubMed: 28984581]

Lezak M, Howieson D, Bigler E, & Tranel D (2012). Neuropsychological Assessment (5th. ed.). New York: Oxford University Press.

Machulda MM, Pankratz VS, Christianson TJ, Ivnik RJ, Mielke MM, Roberts RO, … Petersen RC (2013). Practice effects and longitudinal cognitive change in normal aging vs. incident mild cognitive impairment and dementia in the Mayo Clinic Study of Aging. Clin Neuropsychol, 27(8), 1247–1264. doi:10.1080/13854046.2013.836567 [PubMed: 24041121]

Mathews M, Abner E, Kryscio R, Jicha G, Cooper G, Smith C, … Schmitt FA (2014). Diagnostic accuracy and practice effects in the National Alzheimer's Coordinating Center Uniform Data Set neuropsychological battery. Alzheimers Dement, 10(6), 675–683. doi:10.1016/j.jalz.2013.11.007 [PubMed: 24656850]

McSweeny A, Naugle RI, Chelune GJ, & Luders H (1993). "T-scores for change:" An illustration of a regression approach to depicting change in clinical neuropsychology. The Clinical Neuropsychologist 7, 300–312.

Mormino EC, Betensky RA, Hedden T, Schultz AP, Amariglio RE, Rentz DM, … Sperling RA (2014). Synergistic effect of beta-amyloid and neurodegeneration on cognitive decline in clinically normal individuals. JAMA Neurol, 71(11), 1379–1385. doi:10.1001/jamaneurol.2014.2031 [PubMed: 25222039]

Panza F, Capurso C, D'Introno A, Colacicco AM, Capurso A, Solfrizzi V, & Bennett. (2007). Mild cognitive impairment: risk of Alzheimer disease and rate of cognitive decline. Neurology, 68(12), 964; author reply 964–965. doi:10.1212/01.wnl.0000259691.26674.d3 [PubMed: 17372143]

Panza F, D'Introno A, Colacicco AM, Capurso C, Del Parigi A, Caselli RJ, … Solfrizzi V (2005). Current epidemiology of mild cognitive impairment and other predementia syndromes. Am J Geriatr Psychiatry, 13(8), 633–644. doi:10.1176/appi.ajgp.13.8.633 [PubMed: 16085779]

Petersen RC (2004). Mild cognitive impairment as a diagnostic entity. J Intern Med, 256(3), 183–194. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15324362 [PubMed: 15324362]

Randolph C (2012). Repeatable Battery for the Assessment of Neuropsychological Status. Bloomington, MN: The Psychological Corporation.

Rapport LJ, Axelrod BN, Theisen ME, Brines DB, Kalechstein AD, & Ricker JH (1997). Relationship of IQ to verbal learning and memory: test and retest. J Clin Exp Neuropsychol, 19(5), 655–666. doi:10.1080/01688639708403751 [PubMed: 9408796]

Rapport LJ, Brines D, Axelrod B, & Theisen ME (1997). Full Scale IQ as mediator of practice effects: The rich get richer. Clinical Neuropsychologist, 11(4), 375–380.

Reitan R (1992). Trail Making Test: Manual for administration and scoring. Tucson, AZ: Reitan Neuropsychology Laboratory.

Rentz DM, Locascio JJ, Becker JA, Moran EK, Eng E, Buckner RL, … Johnson KA (2010). Cognition, reserve, and amyloid deposition in normal aging. Ann Neurol, 67(3), 353–364. doi:10.1002/ana.21904 [PubMed: 20373347]

Rinehardt E, Duff K, Schoenberg M, Mattingly M, Bharucha K, & Scott J (2010). Cognitive change on the repeatable battery of neuropsychological status (RBANS) in Parkinson's disease with and without bilateral subthalamic nucleus deep brain stimulation surgery. Clin Neuropsychol, 24(8), 1339–1354. doi:10.1080/13854046.2010.521770 [PubMed: 20967688]

Rodrigue KM, Kennedy KM, Devous MD Sr., Rieck JR, Hebrank AC, Diaz-Arrastia R, … Park DC (2012). beta-Amyloid burden in healthy aging: regional distribution and cognitive consequences. Neurology, 78(6), 387–395. doi:10.1212/WNL.0b013e318245d295 [PubMed: 22302550]

Roe CM, Mintun MA, D'Angelo G, Xiong C, Grant EA, & Morris JC (2008). Alzheimer disease and cognitive reserve: variation of education effect with carbon 11-labeled Pittsburgh Compound B uptake. Arch Neurol, 65(11), 1467–1471. doi:10.1001/archneur.65.11.1467 [PubMed: 19001165]

Roe CM, Xiong C, Grant E, Miller JP, & Morris JC (2008). Education and reported onset of symptoms among individuals with Alzheimer disease. Arch Neurol, 65(1), 108–111. doi:10.1001/archneurol.2007.11 [PubMed: 18195147]

Roh HW, Hong CH, Lee Y, Lee KS, Chang KJ, Kang DR, … Son SJ (2016). Clinical Conversion or Reversion of Mild Cognitive Impairment in Community versus Hospital Based Studies: GDEMCIS (Gwangju Dementia and Mild Cognitive Impairment Study) and CREDOS (Clinical Research Center for Dementia of South Korea). J Alzheimers Dis, 53(2), 463–473. doi:10.3233/JAD-160341 [PubMed: 27163831]

Sanchez-Benavides G, Pena-Casanova J, Casals-Coll M, Gramunt N, Manero RM, Puig-Pijoan A, … Team, N. S. (2016). One-Year Reference Norms of Cognitive Change in Spanish Old Adults: Data from the NEURONORMA Sample. Arch Clin Neuropsychol, 31(4), 378–388. doi:10.1093/arclin/acw018 [PubMed: 27193368]

Schrijnemaekers AM, de Jager CA, Hogervorst E, & Budge MM (2006). Cases with mild cognitive impairment and Alzheimer's disease fail to benefit from repeated exposure to episodic memory tests as compared with controls. J Clin Exp Neuropsychol, 28(3), 438–455. doi:10.1080/13803390590935462 [PubMed: 16618630]

Smith A (1973). Symbol Digit Modalities Test. Los Angeles, CA: Western Psychological Services.

Stein J, Luppa M, Brahler E, Konig HH, & Riedel-Heller SG (2010). The assessment of changes in cognitive functioning: reliable change indices for neuropsychological instruments in the elderly - a systematic review. Dement Geriatr Cogn Disord, 29(3), 275–286. doi:10.1159/000289779 [PubMed: 20375509]

Stern Y (2006). Cognitive reserve and Alzheimer disease. Alzheimer Dis Assoc Disord, 20(2), 112–117. doi:10.1097/01.wad.0000213815.20177.19 [PubMed: 16772747]

Suchy Y, Kraybill ML, & Franchow E (2011). Practice effect and beyond: reaction to novelty as an independent predictor of cognitive decline among older adults. J Int Neuropsychol Soc, 17(1), 101–111. doi:10.1017/S135561771000130X [PubMed: 21073771]

Tabachnick BG, & Fidell LS (1996). Using Multivariate Statistics (Third. ed.). New York: Harper Collins.

Thorgusen SR, Suchy Y, Chelune GJ, & Baucom BR (2016). Neuropsychological Practice Effects in the Context of Cognitive Decline: Contributions from Learning and Task Novelty. J Int Neuropsychol Soc, 22(4), 453–466. doi:10.1017/S1355617715001332 [PubMed: 26790693]

U.S. Bureau of Labor Statistics. (2016). Volunteering in the United States, 2015 [Press release]. Retrieved from https://www.bls.gov/news.release/volun.nr0.htm

United States Bureau of the Census, Statistics, U. S. D. o. L. B. o. L., National, C. f., & Service, C. (2015). Current Population Survey, September 2014: Volunteer Supplement.

Whitwell JL, Wiste HJ, Weigand SD, Rocca WA, Knopman DS, Roberts RO, … Alzheimer Disease Neuroimaging, I. (2012). Comparison of imaging biomarkers in the Alzheimer Disease Neuroimaging Initiative and the Mayo Clinic Study of Aging. Arch Neurol, 69(5), 614–622. doi:10.1001/archneurol.2011.3029 [PubMed: 22782510]

Wilkinson GS, & Robertson GJ (2006). WRAT 4: Wide Range Achievement Test, professional manual. Lutz, FL: Psychological Assessment Resources, Inc.

Yan JH, & Dick MB (2006). Practice effects on motor control in healthy seniors and patients with mild cognitive impairment and Alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn, 13(3–4), 385–410. doi:10.1080/138255890969609 [PubMed: 16887780]

Table 1.

Demographic characteristics of Hammers et al.'s (2020) development and the current amnestic MCI validation samples

Variable	Current Validation Sample		Hammers et al., 2020 Development Sample
	Community Amnestic MCI	Clinic Amnestic MCI	
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)
n	64	58	129
Age (years)[1, 2]	79.2 (7.9)	74.1 (5.2)	75.6 (7.5)
Education (years)[1, 3]	15.6 (2.6)	17.0 (2.9)	15.4 (2.7)
Gender (*n; %*)[1, 2, 3]			
Males	21 (33%)	38 (66%)	22 (17%)
Females	43 (67%)	20 (36%)	107 (83%)
Ethnicity (*n*)			
Non-White, Non-Caucasian	0	1	1
White, Non-Hispanic	64	57	128
WRAT-4 Reading Subtest	109.8 (10.8)	107.2 (7.9)	107.4 (7.6)
Baseline RBANS Indexes			
Immediate Memory[1, 2, 3]	86.6 (13.6)	76.9 (14.4)	109.2 (14.5)
Visuospatial/Constructional[3]	100.4 (14.9)	98.8 (11.4)	105.2 (15.7)
Language[1, 2, 3]	98.3 (11.4)	88.6 (10.5)	104.6 (11.2)
Attention[2, 3]	95.9 (14.1)	96.7 (14.9)	104.4 (14.7)
Delayed Memory[1, 2, 3]	80.2 (12.7)	69.7 (19.2)	108.6 (9.1)
Total Scale[1, 2, 3]	89.2 (10.4)	81.6 (11.3)	109.2 (12.7)
Retest Interval (days)[1, 3]	398.5 (66.8)	474.2 (49.5)	381.5 (37.9)

Note: MCI = Mild Cognitive Impairment, WRAT-4 = Wide Range Achievement Test,

1 = significant difference between the current validation community amnestic MCI and clinic amnestic MCI samples, $p < .01$.

2 = significant difference between the current validation community amnestic MCI sample and Hammers et al.'s development sample, $p < .01$.

3 = significant difference between the current validation clinic amnestic MCI sample and Hammers et al.'s development sample, $p < .01$.

Table 2.

Regression equations for predicting Time 2 RBANS Indexes from Hammers (2020) development sample.

RBANS Index Scores	Predicted T_2	R^2	SE_{est}
Immediate Memory	78.98 + (T_1*0.59) − (*age**.37)	.42	10.89
Visuospatial/Constructional	56.85 + (T_1*0.30) + (*age**.12) + (*ed**.67) − (*sex**7.41)	.17	14.03
Language	40.26 + (T_1*0.60) − (*age**.11) + (*ed**.74)	.36	9.84
Attention	39.76 + (T_1*0.63)	.34	13.09
Delayed Memory	34.91 + (T_1*0.70)	.25	10.86
Total Scale	31.32 + (T_1*0.72)	.38	11.79

Note: Index scores are age-corrected standardized scores, subtest scores are raw scores. Age and education (*ed*) are in years. Refer to Method section for coding of ethnicity and gender. R^2 = squared value of Pearson's correlation coefficient for initial and retest score, SE_{est} = Standard error of the estimate. To calculate the Predicted Time 2 (T_2) score, use the formula in the column titled "Predicted T_2". To calculate the reliable change score, use (Observed T_2 – Predicted T_2) / SE_{est}.

Table 3.

Baseline, Follow-up, and Predicted Follow-up RBANS scores, standardized *z* scores, and *p* values for Observed versus Predicted Follow-up scores in community amnestic MCI participants (*n* = 64)

	Observed Baseline	Observed Follow-up	Predicted Follow-up	*z* score	*p* Value
RBANS Indexes					
Immediate Memory	86.6 (13.6)	92.5 (13.9)	96.8 (8.4)	−0.40 (1.1)	0.007
Visuospatial/Constructional	100.4 (14.9)	96.7 (16.8)	101.9 (7.0)	−0.37 (1.1)	0.008
Language	98.3 (11.4)	96.0 (11.2)	102.1 (7.1)	−0.62 (1.0)	0.001
Attention	95.9 (14.1)	97.8 (15.4)	100.2 (8.9)	−0.19 (0.9)	0.09
Delayed Memory	80.2 (12.7)	87.0 (18.4)	91.0 (8.9)	−0.38 (1.2)	0.02
Total Scale	89.2 (10.4)	91.8 (12.8)	95.5 (7.5)	−0.32 (0.8)	0.003

Note: RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, *p* value = significance of one-sample *t* tests examining whether *z* scores differed from expectation (z = 0) based on the normal distribution of *z* scores.

Table 4.

Baseline, Follow-up, and Predicted Follow-up RBANS scores, standardized *z* scores, and *p* values for Observed versus Predicted Follow-up scores in clinic amnestic MCI participants ($n = 58$)

	Observed Baseline	Observed Follow-up	Predicted Follow-up	*z* score	*p* Value
RBANS Indexes					
Immediate Memory	76.9 (14.4)	76.0 (16.3)	93.3 (8.3)	−1.60 (1.2)	0.001
Visuospatial/Constructional	98.8 (11.4)	99.0 (16.5)	104.2 (7.3)	−0.37 (0.9)	0.004
Language	88.6 (10.5)	86.2 (12.2)	97.9 (6.5)	−1.19 (1.1)	0.001
Attention	96.7 (14.9)	93.5 (15.9)	100.7 (9.4)	−0.55 (0.9)	0.001
Delayed Memory	69.7 (19.2)	68.6 (20.6)	83.7 (13.5)	−1.39 (1.3)	0.001
Total Scale	81.6 (11.3)	80.2 (12.8)	90.1 (8.1)	−0.84 (0.7)	0.001

Note: RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, *p* value = significance of one-sample *t* tests examining whether z scores differed from expectation ($z = 0$) based on the normal distribution of *z* scores.

Table 5.

Percentage of community and clinic amnestic MCI samples that declined, remained stable, or improved based on standardized regression-based methodology (total $n = 122$)

	Community Amnestic MCI ($n = 64$)			**Clinic Amnestic MCI ($n = 58$)**		
	Decline	**Stable**	**Improve**	**Decline**	**Stable**	**Improve**
RBANS Indexes						
Immediate Memory	11	84	5	47	53	0
Visuospatial/ Constructional	11	84	5	9	90	1
Language	13	84	3	17	83	0
Attention	5	92	3	9	90	1
Delayed Memory	16	80	5	47	52	1
Total Scale	3	95	2	17	83	0
Cumulative Percent of Decline/Stable/Improve Across Indexes	10	86	4	24	75	1

Note: MCI = Mild Cognitive Impairment, RBANS = Repeatable Battery for the Assessment of Neuropsychological Status.

Table 6.

Case Example

	Observed Baseline	Observed Follow-up	Observed Difference	Predicted Follow-up	Predicted Difference	SE_{est}	*z* score
RBANS Indexes							
Immediate Memory	81	79	−2	97.37	−18.37	10.89	−1.69*
Visuospatial/Constructional	95	100	+5	101.79	−1.79	14.03	−0.13
Language	95	110	+15	98.44	11.56	9.84	1.17
Attention	88	85	−3	95.20	−10.20	13.09	−0.78
Delayed Memory	78	71	−7	89.51	−18.51	10.86	−1.70*
Total Scale	83	85	+2	91.08	−6.08	11.79	−0.52

Note: Observed Difference = Observed Follow-up – Observed Baseline. Predicted Follow-up scores are derived from the regression formula from Hammers et al. (2020). Predicted Difference = Observed Follow-up – Predicted Follow-up. SE_{est} = Standard Error of the Estimate of the regression equations from Hammers et al. (2020). *z* = Predicted Difference/ SE_{est}.

* $p < .05$.