HHS Public Access

Author manuscript

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

Published in final edited form as:

Clin Neuropsychol. 2022 November ; 36(8): 2221–2236. doi:10.1080/13854046.2021.1952308.

Demographically-corrected normative data for the RBANS Learning Ratio in a sample of older adults

Dustin B. Hammers[1], **Kevin Duff**[2], **Robert J. Spencer**[3,4]

[1]Department of Neurology, Indiana University School of Medicine, Indianapolis, IN, USA

[2]Center for Alzheimer's Care, Imaging, and Research, Department of Neurology, University of Utah, Salt Lake City, UT, USA

[3]Mental Health Service, VA Ann Arbor Healthcare System, Ann Arbor MI, USA

[4]Michigan Medicine, Department of Psychiatry, Neuropsychology Section, Ann Arbor MI, USA

Abstract

Background: A novel learning slope score – the Learning Ratio (LR) – has recently been developed that appears to be sensitive to memory performance and AD pathology more optimally than traditional learning slope calculations. While promising, this research to date has been both experimental and based on group differences, and therefore does not aid in the interpretation of individual LR performance for either clinical or research settings. The objective of the current study was to develop demographically-corrected normative data on these LR learning slopes on verbal learning measures from the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS).

Method: The current study examined the influence of age and education on LR metrics for the List Learning, Story Memory, and an Aggregated RBANS score in 200 cognitively intact adults aged 65 or older using linear regression.

Results: Age and education correlated with most LR metrics, but no sex differences were observed. Linear regression permitted the prediction of LR values from age and education, which are then compared to observed LR values. The result is demographically-corrected *T* scores for these LR metrics.

Conclusions: By comparing observed and predicted LR scores calculated from regression-based prediction equations, this represents the first step towards interpretation of individual performances on this metric for clinical decision making and treatment planning purposes. With future replication in diverse and heterogenous samples, we hope to offer a new clinical tool for the examination of learning slopes in older adults.

Keywords

Learning; Memory; Alzheimer's disease; Mild Cognitive Impairment; RBANS

Corresponding author: Dustin B. Hammers, Ph.D., ABPP(CN), Indiana University School of Medicine, 355 West 16th Street (GH4027), Indianapolis, IN, 46202, hammersd@iu.edu.

INTRODUCTION

Neuropsychological evaluations of older adults traditionally assess learning and retention of information over multiple trials (Lezak et al., 2012; Suhr, 2015). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS;Randolph, 2012) is an example of a cognitive measure that has been frequently administered since its creation two decades ago and contains relevant learning data. Specifically, the List Learning and Story Memory subtests of the RBANS assess learning and immediate memory, and can be used in conjunction with their delayed recall counterparts (RBANS subtests List Recall, List Recognition, Story Recall, and Figure Recall) to evaluate encoding and retention over time. Additionally, the steepness – or pitch – of the learning slope can be informative in terms of the potential for an individual to benefit from repeated exposure to stimuli over multiple trials. Learning slopes are often shallow in individuals with learning and memory impairments, including in conditions like Alzheimer's disease (AD; Gifford et al., 2015), Frontotemporal Dementia (Lemos et al., 2014), Vascular Dementia (Mast & Allaire, 2006), and Attention Deficit Hyperactivity Disorder (ADHD; Skodzik et al., 2017). Although many test manuals provide basic information about learning slope data, the calculations for such learning slopes tend to be relatively simplistic – traditionally only considering the difference between the final and first learning trial ("Final Trial minus Trial One"). Consequently, potential exists for these data to provide more a nuanced understanding of learning in some patients.

In contrast to the traditional approach, an alternate metric for calculating learning slope has been developed. Spencer and colleagues (Spencer et al., 2020) recently developed the Learning Ratio (LR) – which is a learning slope metric that examines the proportion of information learned over successive trials relative to the information available to learn. Spencer et al.'s LR metric is calculated as the number of items learned on subsequent trials after Trial One, divided by the number of unlearned items on Trial One. As a result, this metric considers the opportunity for future learning, which varies as a function of Trial One success. Let us suppose a patient learns 2 items on Trial One of a 10-item word list and 7 items on the final trial. This individual would obtain an LR value of 0.63, as the patient learned 5 items over successive trials after Trial One out of a potential of 8 additional items to be learned after Trial One (5/8 = 0.63). In dividing by the number of yet-to-be-learned items, LR controls for performance on initial learning trials to better enable a more accurate representation of learning capacity. As LR is a proportion of information learned over successive trials relative to information available to learn, our example patient's LR of 0.63 means that the patient learned 63% of information still left to learn after Trial One. This is novel because in traditional "Final Trial minus Trial One" learning slope calculations, there exists an inherently inverse relationship between Trial One results and the amount of information remaining to learn; in essence, the more information an individual learns at Trial One, the less information is available in which to improve over successive trials. However, this competition between Trial One and subsequent trials is accounted for with Spencer et al.'s LR calculation, thus eliminating the confound.

A handful of studies highlight the incremental improvement of this LR metric over traditional "Final Trial minus Trial One" calculations – and the overall convergent and

criterion validity for this LR metric. Spencer et al. (2020) originally developed and subsequently validated LR scores from the immediate memory subtests of RBANS in 289 older veteran patients from an outpatient neuropsychology clinic. In Spencer et al.'s study, LR equations for List Learning, Story Memory, and an Aggregated RBANS metric exhibited stronger correlations with common memory measures than traditional learning slope calculations, and were superior in discriminating between individuals with and without a neurocognitive diagnosis (Spencer et al., 2020). Hammers et al. (In Press) additionally validated Spencer et al.'s LR calculation applied to the Hopkins Verbal Learning Test – Revised (HVLT-R; Brandt & Benedict, 2001) and the Brief Visual Memory test – Revised (BVMT-R; Benedict, 1997) in an independent sample of 56 clinical patients with a cognitive disorder. They observed that smaller HVLT-R LR, BVMT-R LR, and Aggregated HVLT-R/BVMT-R LR values corresponded to smaller total hippocampal volumes and worse performances on standard memory measures. Patients with AD additionally achieved smaller LR values than those with Mild Cognitive Impairment (MCI). Similar to the Spencer and colleagues' study, the magnitude of the effects observed in Hammers et al. were consistently larger for this LR metric than the traditional "Final Trial minus First Trial" learning-curve calculation (Hammers et al., In Press). In a separate study from a sample of patients across the AD continuum, Hammers and colleagues (Hammers, Suhrie, Dixon, Gradwohl, Duff, et al., 2021) have also found that patients with MCI and AD perform worse on this LR metric than a cognitively intact group for the RBANS List Learning, RBANS Story Memory, HVLT-R, BVMT-R, and Aggregate LR values. LR scores also displayed strong receiver operator characteristics when differentiating individuals with and without cognitive impairment, and lower LR values were associated with lower performances on traditional memory measures after controlling for the demographic variables of age and education.Further, LR values from the RBANS List Learning, RBANS Story Memory, HVLT-R, and BVMT-R have additionally shown criterion validity with AD biomarkers of amyloid burden, hippocampal volumes, and apolipoprotein e4 carrier status (Hammers, Suhrie, Dixon, Gradwohl, Archibald, et al., 2021).

These recent studies suggest that LR is adept at characterizing learning capacity, and is also sensitive to performance along the AD continuum. However, all LR results thus far are experimental and based on group-level analyses, therefore they cannot be directly applied to understand individual performance in either clinical or research settings. For this LR metric to maximize clinical utility, we offer normative data for LR values from common memory measures in cognitively intact individuals(Goodwill et al., 2019; Harrington et al., 2017). Normative data correcting for age and education have recently been developed for LR values derived from the HVLT-R and BVMT-R (Hammers, Duff, et al., 2021), consequently the purpose of the current study is to develop demographically-corrected normative data for the LR metric when applied to the List Learning and Story Memory subtests of the RBANS – and an Aggregated LR metric for the two subtests – in a sample of older adults with intact cognition. As research has previously shown relationships between common demographic variables (e.g., age, education, and sex) and (1) the RBANS (Duff et al., 2003; Duff & Ramezani, 2015; Duff et al., 2011), (2), other standard neuropsychological measures ("MOANS norms"; Ivnik et al., 1996; Ivnik et al., 1992), (3) and LR values for the HVLT-R and BVMT-R (Hammers, Duff, et al., 2021), it was hypothesized that these variables would

be both associated with and predictive of the LR metrics calculated from the verbal memory subtests of the RBANS in this older adult sample. By creating a normative sample for Spencer et al.'s LR metric, we aim to bolster its use as a clinical tool for the examination of verbal learning slopes in older adults who are administered the RBANS.

METHODS

Sample and Procedure

Participants were drawn from two prior studies that evaluated MCI and practice effects, and AD biomarkers and practice effects. Both samples were recruited from senior centers and independent living facilities. Specifically, Sample One was comprised of 148 cognitively intact community-dwelling older adults recruited as a control group for a study of practice effects and MCI from 2008 to 2013 (see Duff et al., 2017). Sample Two included 52 cognitively intact community-dwelling older adults recruited from 2019 to present as a control group for a study of practice effects and AD biomarkers(Hammers, Suhrie, Dixon, Gradwohl, Duff, et al., 2021). Of note, these 52 participants were also used as the Normal Control group for two aforementioned studies examining LR's relationship with standard memory measures (Hammers, Suhrie, Dixon, Gradwohl, Duff, et al., 2021) and AD biomarkers (Hammers, Suhrie, Dixon, Gradwohl, Archibald, et al., 2021) along the AD continuum. The mean age of Sample One was 75.6 (*SD* = 7.1, range = 65 – 96) years old, and the mean age of Sample Two was 72.5 (*SD* = 4.9, range = 65 – 91) years old. Slightly over half of the participants (54%) were between the ages of 65 and 74, 37% of participants were between the ages of 75 and 84, and 9% of participants were aged 85 or above. Sample One averaged 15.4 (*SD* = 2.6, range = 8 – 20+) years of education, and Sample Two averaged 16.7 (*SD* = 2.1, range = 12 – 20) years of education. Both samples were predominantly Caucasian, with Sample One being predominantly female (83.1% female) and Sample Two possessing a marginally higher proportion of females than males (61.5% female). Premorbid intellect at baseline for both samples was estimated to be average to high average according to the Wide Range Achievement Test – Third and Fourth editions (WRAT; Wilkinson, 1993; Wilkinson & Robertson, 2006) Reading subtest (*standard score*: *M* = 107.4, *SD* = 7.2, range = 81 – 126 for Sample One, and *M* = 110.6, *SD* = 7.4, range = 88 – 126 for Sample Two). Self-reported depression for both samples was low, including a mean of 4.0 (*SD* = 3.6, range = 0 – 14) according to the 30-item Geriatric Depression Scale (GDS; Yesavage et al., 1982) for Sample One, and a mean of 0.9 (*SD* = 1.0, range = 0 – 5) for Sample Two using the 15-item GDS (Sheikh & Yesavage, 1986). It should be noted that self-reported depression was part of the exclusion criteria for the parent study of Sample Two, therefore scores were not observed of GDS ≥ 5 in this sample.

For inclusion in the present study, participants from both samples were classified as being cognitively intact/free of cognitive impairment (e.g., MCI or dementia due to AD). Classification of participants from Sample One has been described previously (Duff et al., 2017). Briefly, all participants in this sample performed within 1.5 SD of the mean for each domain of a baseline cognitive evaluation described below – which included measures that were separate from RBANS performance to avoid diagnostic circularity. Classification of participants from Sample Two was based on the classification battery developed in the

Alzheimer's Disease Neuroimaging Initiative (ADNI2, 2020), which included the Mini Mental State Examination (Folstein et al., 1975), the Wechsler Memory Scale-Revised (Wechsler, 1987) Logical Memory II Paragraph A, and the Clinical Dementia Rating Scale (Morris, 1993).

The two cognitively intact samples differed on age, $t(130.19) = 3.49$, $p = .001$, $d = 0.56$, education, $t(111.11) = -3.09$, $p = .001$, $d = -0.50$, sex, χ^2 (1) = 9.07, $p = .003$, *Phi* = −0.23, and premorbid intellect, $t(198) = -2.76$, $p = .008$, $d = -0.44$. No sample differences were observed for ethnic distribution, χ^2 (1) = 0.01, $p = .99$, *Phi* = −0.02. Although significant, when viewing the group means, any differences between samples were generally small in magnitude, and reflected normal variation within the distribution of intact individuals. As a result, the two samples were pooled together to create a cognitively intact combined normative sample with a total sample size of 200 participants. Please see Table 1 for demographic information for the combined normative sample, which generally displayed within to above average abilities for immediate and delayed memory skills, executive functioning, attention, visuospatial skills, and language.

General inclusion criteria for the current study involved being age 65 years or older and functionally independent (according to participant and/or knowledgeable informant). Possession of adequate vision, hearing, and motor abilities to complete the cognitive evaluation was also necessary. Additional exclusion criteria included neurological conditions likely to affect cognition, dementia, major psychiatric condition, current severe depression, substance abuse, anti-convulsant or anti-psychotic medications, or residence in a skilled nursing or living facility. Further, as described above, the second sample used a cut-off of 5/15 (or higher) on the 15-item Geriatric Depression Scale (Sheikh & Yesavage, 1986) as exclusion for the parent study.

Procedure—All procedures were approved by a local Institutional Review Board prior to the initiation of the study. All participants provided informed consent before completing any procedures. The following primary measures were administered:

- The RBANS (Randolph, 2012) 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. The two subtests germane to the current analyses include List Learning (four learning trials of ten items) and Story Memory (two learning trials of twelve items). Of note, the index scores utilize age-corrected normative comparisons from the test manual to generate *standard scores* ($M = 100$, $SD = 15$). Individual subtest scores for List Learning and Story Memory use age-corrected normative comparisons generated *Tscore* values ($M = 50$, $SD = 10$). Learning slope performances were evaluated by raw data from individual trials of each of the memory measures. For raw scores and all normed comparisons (*standard scores* and *T scores)*, higher values indicate better performance.

- WRAT Reading subtest – Third and Fourth editions (Wilkinson, 1993; Wilkinson & Robertson, 2006) are used as an estimate of premorbid intellect for the first and second samples, respectively. During this task an individual attempts to pronounce words, and the

raw score is normalized to standard scores ($M = 100$, $SD = 15$) relative to age-matched peers. Higher values indicate better performance.

Additional measures of cognition were administered in the parent studies for diagnostic classification purposes (for Sample One) and are included in the tables to further cognitively describe the sample. HVLT-R (Brandt & Benedict, 2001) is a verbal memory task with 12 words learned over three trials, with the correct words summed for the Total Recall score (range = 0 – 36). The Delayed Recall score is the number of correct words recalled after a 20 – 25-minute delay (range = 0 – 12). BVMT-R (Benedict, 1997) is a visual memory task with 6 geometric designs in 6 locations on a card learned over three trials, with correct designs and locations summed for the Total Recall score (range = 0 – 36). The Delayed Recall score is the number of correct designs and locations recalled after a 20 – 25-minute delay (range = 0 – 12). The Symbol Digit Modalities Test (Smith, 1973) is a divided attention and psychomotor speed task, with the number of correct responses in 90 seconds being the total score (range = 0 – 110). Finally, Trail Making Test, Parts A and B (Reitan, 1992) are tests of visual scanning/processing speed and set shifting/complex mental flexibility, respectively. For each subtest, the score is the time to complete the task (range = 0 – 180 seconds for Part A, and range = 0 – 300 seconds for Part B). Age-corrected normative comparisons were used for descriptive statistics for the memory measures HVLT-R and BVMT-R, and age- and education-corrected normative comparisons were used for Symbol Digit Modalities Test and the Trail Making Test, Parts A and B; for all four measures, normative comparisons generated *T score* values ($M = 50$, $SD = 3$), with higher scores indicating better cognition.

Calculation of Learning Slopes

For List Learning, Story Memory, and the Aggregated RBANS memory measures, the LR score was computed as a proportion where differences in performance between the Final Trial and Trial One is in the numerator, and the difference between a maximum score for a trial and performance on Trial One serves as the denominator. This calculation is consistent with previous research (Hammers et al., In Press; Spencer et al., 2020). The aggregated LR score was computed as the combined difference between Trial One and the Final Trial for both tests, divided by the difference between the combined total points available for a trial for both tests and the sum of Trial One from both tests. The formulas for LR for the List Learning, Story Memory, and the Aggregated RBANS are as follows:

$$\text{LR for each test} = \frac{(\text{Final Trial} - \text{Trial One})}{(\text{Total Points Available For a Single Trial} - \text{Trial One})}$$

$$\text{Aggregated LR} = \frac{(\text{Measure 1: Final Trial} - \text{Trial One}) + (\text{Measure 2: Final Trial} - \text{Trial One})}{(\text{Combined Total Points Available for a Trial from Both Tests} - \text{Sum of Trial One from Both Tests})}$$

Data Analysis

Independent samples *t* tests for the continuous demographic variables (e.g., age, education, and premorbid intellect) and chi square analyses for the dichotomous demographic variables (e.g., sex and ethnicity) were calculated to assess the appropriateness of combining the two

samples into a larger normative sample. Bivariate correlation coefficients were subsequently calculated between LR values and demographic variables of age and education in the pooled normative sample to examine their influence on the LR metrics. Independent samples *t* tests were also calculated for the categorical demographic variable of sex for List Learning LR, Story Memory LR, and the Aggregated RBANS LR in the pooled sample.

To generate demographically-corrected normative data, linear regression analyses were conducted to predict the List Learning LR, Story Memory LR, and Aggregated RBANS LR values (Cherner et al., 2007; Duff, 2016; Norman et al., 2011). Specifically, demographic variables of age and education were the predictor variables, and the individual LR values were the criterion variable in each regression analysis. Sex was not included in the model because descriptive analyses did not show an association between sex and RBANS LR performance.

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 .05 for all statistical analyses.

RESULTS

Demographics and Memory Testing

The demographic characteristics of the current normative sample can be observed in Table 1, along with the sample's performance among other neuropsychological domains. Table 2 displays the sample's mean and *SD* values for the List Learning, Story Memory, and the Aggregated RBANS LR metrics, along with individual trial performances for the RBANS memory measures. The mean List Learning LR value for the 200 participants in the current study was 0.704 (*SD* = 0.258; range 0.00 – 1.00), the mean Story Memory LR value was 0.710 (*SD* = 0.296; range −1.00 – 1.00), and the mean Aggregated RBANS LR value was 0.710 (*SD* = 0.200; range 0.13 – 1.00). This equates to the sample, on average, learning 70% to 71% of the available information after Trial One for these memory measures. Table 2 also indicates that bivariate correlation coefficients were significant between age and both List Learning LR and Aggregated RBANS LR (*p*s< .005), and that there was a significant correlation between education and both Story Memory LR and Aggregated RBANS LR metrics (*p*s< .02). Similarly, individual trials of List Learning performance were significantly correlated with age (*p*s< .001) and education (*p*s = .001 to .03), and individual trials of Story Memory performance were correlated with education (*p*s = .004 to .05) and age (for Trial One; *p* = .001). In contrast, no sex differences were observed for any LR metric – including for List Learning LR, $t(198) = -1.08$, $p = .28$, $d = -0.15$, Story Memory LR, $t(198) = 0.24$, $p = .81$, $d = 0.03$, nor the Aggregated RBANS LR, $t(198) = -1.03$, $p = .30$, $d = -0.15$. No sex differences were observed for most List Learning or Story Memory individual trials (*p*s = .05 to .99; the only exception being for List Learning Trial 2 at *p* = .046). As a result, demographic variables of age and education were included in the subsequent linear regression analyses.

Linear Regression Analyses

Linear regression analyses were conducted in the current study for the List Learning LR, Story Memory LR, and Aggregated RBANS LR as the criterion variables and age and education as the predictor variables. These results can be observed in Table 3. Briefly, the model containing both demographic variables significantly predicted all three LR metrics (*p*s = .001 to .04).

DISCUSSION

The current study aimed to develop demographically-corrected normative data in older adults for the Learning Ratio (LR), which is a recently created and validated metric for assessing learning slope that controls for initial trial learning. When applying the LR to the List Learning and Story Memory measures of the RBANS, results suggested that the demographic variables age and education were significantly correlated with RBANS LR performance in our current sample, but sex was not (see Table 2). This is generally consistent with our expectations. Specifically, bivariate correlations indicated that two of the three LR calculations – List Learning LR and Aggregated RBANS LR – were negatively correlated with age, such that increased age was associated with worse LR performance. Similarly, education was positively correlated with two of the three LR metrics – Story Memory and Aggregated RBANS LR – with greater levels of education being associated with greater LR performance. However, the lack of association between List Learning LR and education – and the generally smaller associations for education relative to age – were surprising. It is possible that the generally stronger overall associations between LR values and age (as compared to education) were related to the relative restriction of range of education in our sample, as restricted ranges are known to result in smaller correlation coefficients (Bland & Altman, 2011). These LR results are similar to those observed between demographic variables and the individual learning trials for the List Learning and the Story Memory subtests in the current study. They additionally coincide with research suggesting associations between age/education and performance on both memory and non-memory subtests in the RBANS (Duff et al., 2003; Duff & Ramezani, 2015), though it is of interest that our study failed to find an association between these RBANS LR metrics and sex. This is in contrast to studies by Duff and Colleagues (Duff & Ramezani, 2015; Duff et al., 2011) showing that List Learning and Story Memory were associated with sex, though differences in sample size (200 currently versus 718 for both (Duff et al., 2011) and Duff and Ramezani(2015), sex distribution (78% female currently versus 58% for both Duff studies), and diagnostic make-up (cognitively intact participants currently versus a mixed clinical sample in both Duff studies) may have led to these differential outcomes. A study by Gogos and colleagues (Gogos et al., 2010) similarly failed to find a difference in RBANS performance based on sex. Despite this lack of association for sex, overall these current data support the use of demographic corrections in the normative data for the RBANS verbal memory LR metrics.

As can be observed in Table 3, regression-based prediction equations were developed for the List Learning LR, Story Memory LR, and Aggregated RBANS LR. For each LR metric, predicted LR values were generated from models containing the demographic variables

of age and education. Upon closer inspection of the R^2 values, the models accounted for a relatively small proportion of variance (7%, 3%, and 6% for List Learning LR, Story Memory LR, and Aggregated RBANS LR, respectively). The small bivariate correlations with LR for age and education in Table 2 suggest that this finding is not surprising, which also corresponds to the smaller beta weights for age and education in the prediction equations in Table 3. They also coincide with data from Duff and Ramezani(2015) indicating that models containing age, education, sex, and race collectively only accounted for 12% of the variance with List Learning, and 13% of the variance with Story Memory. Despite the lower magnitude of variance accounted for, age and education significantly predicted each of our LR metrics, and their inclusion permits more accurate and patient-specific normative comparisons than using psychometric conversion alone (with LR mean and *SD*). Additionally, while there may not be consensus in neuropsychology about the need for these additional corrections, it has been suggested that additional contributions of age and education could be useful for certain patients, especially those at the extremes (e.g., very old, very high or low levels of education; Duff, 2016).

Table 4 displays an example of how to apply these prediction equations to LR performance for an individual, though the interested reader can also contact the first author to obtain an Excel spreadsheet that will automatically calculate these demographically-corrected values. Overall, these prediction equations generate predicted LR values for each RBANS verbal memory subtest, which can be compared to the observed LR values to assess individual deviation from an individual's same-aged and -educated peers. The specific details – including an example calculation – are listed here. After calculating the observed LR value for the individual using the equations in the Methods, the discrepancy between the observed LR value and the predicted LR value is calculated (observed LR – predicted LR/ Standard Error of the Estimate [SE_{est}]). For the example of a 69-year-old woman with 15 years of education who obtained an observed List Learning LR value of 0.86 (recalling 86% of available information after Trial One), her predicted List Learning LR value was 0.78. The discrepancy between the observed and predicted LR values was 0.08, and when divided by 0.25 (the SE_{est} from Tables 3 or 4) led to a z value of 0.34. Lastly, this LR Discrepancy z-score value is then translated into an age- and education-corrected T score (multiplying by 10, adding 50). Consequently, a List Learning LR score of 0.86 is equivalent to a T score of 54 for a 69-year-old woman with 15 years of education, which corresponds to performance near the upper limit of the average range.

Although the creation of look-up tables for each age and education group would result in dozens of tables, and is beyond the scope of this manuscript, Table 5 reflects the performance distribution on the LR metrics for a 75-year-old with 16 years of education. As can be observed, for this set of demographic characteristics the mean (T = 50) raw LR values would be approximately 0.70 for each of List Learning, Story Memory, and the Aggregated RBANS, suggesting that the average 75-year-old with 16 years of education learned 70% of available information after Trial One on subsequent trials. This also corresponds to the LR means in Table 2. Weaker performances (T< 44) on the LR metrics begin to be observed between LR values of 0.55 to 0.60 (55% to 60% of available information learned), and impaired performances (T< 31) on the LR metrics were below 0.18–0.23 (18–23% learned) for List Learning and Story Memory, and below 0.35 (35% learned) for the Aggregated

LR metric. A closer inspection of Table 5 also shows that the regression equations appear to result in a ceiling effect for the RBANS LR performances. Specifically, an individual can perform quite poorly using these norms (learning 0% results in a *T* score of 21 for List Learning LR and 25 for Story Memory LR), but performance beyond the high average range is not possible (a learning score of 100% results in a *T* score of 62 for List Learning LR and 59 for Story Memory LR). As such, the metric appears to be slightly more sensitive at identifying individuals with learning problems compared to discriminating between individuals with intact versus exceedingly strong learning capacities.

Further clinical benefit of the use of this LR metric can be observed with a finer examination of the individual performance data in our example of a 69-year-old woman with 15 years of education. Specifically, by applying this normative data at the level of the individual, we can now possess a more nuanced understanding of learning for some patients. In our example (Table 4), her List Learning Total Recall performance was $T = 43$, her Story Memory Total Recall performance was $T = 37$, and her aggregated Total Recall performance (i.e., RBANS Immediate Memory Index converted to a *T* score) was $T = 39$, according to the test manuals. This equates to learning abilities consistently being in the low average range for this individual. However, when examining her individual trial performances on the RBANS verbal memory subtests, it can be seen that she tends to possess a steep learning curve, with limited initial trial learning but improvement with successive exposures to the material. Her observed LR values of 0.86, 0.75, and 0.80 (for List Learning, Story Memory, and the Aggregated RBANS, respectively) suggest that she learned 75% to 86% of available information after Trial One on subsequent trials; after applying the current demographically-corrected normative comparisons, these performances equate to *T* scores of 54, 51, and 52, respectively. These values are consistent with the middle to the upper limit of the average range, and imply greater learning capability for this individual than her Total Recall *T* scores would indicate. In essence, this data suggests a clinical picture of an individual with a strong capacity to benefit from repetition, though also with weaker learning upon initial exposure. As a result, treatment recommendations could highlight the particular importance for this individual of repetition and multiple exposure to incoming information. Consequently, these norms are felt to provide a more detailed clinical picture of learning performance, and can aid in treatment recommendations for some patients.

The current study is not without limitations. First, the sample in the current study was homogenous with regards to premorbid functioning, education, ethnicity, and sex, such that few participants in the current sample were non-Caucasian, and the sample was predominantly female. The current study's demographics appear to reflect long-standing trends in research participation; research suggests that women tend to volunteer more than men across all age ranges (upwards of 30% higher rates; (United States Bureau of Labor Statistics, 2016), and that individuals with higher education and Caucasians consistently volunteer at greater levels (United States Bureau of the Census, 2015). Regardless of the similarities with historical trends, however, this represents an important limitation for this study. It is unclear how these normative comparisons perform in populations of other ethnicities. While no sex-based differences were observed for any LR metrics in the current study, it is possible that our results would have differed if our sample was more evenly represented across sex. As a consequence, it will be essential for future research to

replicate these findings with a more diverse and heterogeneous sample before their being used clinically. Second, as all individuals in our sample were 65 years or older, these demographically-corrected normative comparisons do not provide information for younger individuals. Relatedly, a relatively small number of our participants were over the age of 90 (n = 4), consequently future studies with larger numbers of participants aged 90+ are encouraged to examine whether "super-agers" perform comparably or discrepantly with 80 year-old participants on this LR variable. Third, data on the lack of diagnoses of ADHD or learning disorder were available for only Sample Two (n = 52), consequently it is possible that a portion of the participants from Sample One (n = 148) possessed previous diagnoses of ADHD or learning disorder; that said, all participants in this study were classified as being cognitively intact based on objective testing, consequently it is less likely that ADHD or learning disorder would have significantly impacted learning abilities in participants from Sample One. Fourth, these results are specific to participants having been exposed to other memory measures (e.g., HVLT-R, BVMT-R, or Logical Memory), therefore the participants are not naïve to cognitive testing. Although this exposure to other memory measures may have influenced the current results, it was necessary to ensure that participants were cognitively intact while avoiding diagnostic circularity. Fifth, these results were also specific to the LR metric derived from the List Learning and Story Memory subtests from the RBANS, using the equations from Spencer et al. (2020). Additional research is encouraged to consider normative data for LR metrics from other memory measures, particularly those with a visual learning component. Further, an LR value of 1.00 (100% learned) was assigned in the rare scenario when participants learned all available items on Trial One of either RBANS measure; this was consistent with Spencer et al., and necessary because a "perfect" score on Trial One would result in a value of zero in the denominator (Total Points Available For a Single Trial – Trial One), which represents an undefined mathematical expression. Seventh, individual trial stability has been found to be consistently low across memory measures, which reflects a limitation of all learning slope research. This is exacerbated by the restriction of range commonly observed in individual trial performance across test manuals when examining cognitively healthy controls. As a result, caution should be exercised when interpreting these normative values, and we would encourage the use of broader confidence intervals when using these metrics in clinical practice. Finally, future research should consider additional demographic information (beyond age and education) to potentially improve prediction accuracy of these regression equations, though it is of note that these demographic variables were selected based on ease of availability and their significant relationships with the criterion variables.

Despite these limitations – in particular our use of a rather homogenous normative sample – the current study is the first to calculate demographically-corrected normative comparisons for the LR learning slope metric for the verbal memory subtests of the RBANS. This represents the first step towards translating this experimental measure to the clinic. Future steps should include further research to examine the validity of these normative comparisons in more diverse samples, which will aid interpretation of individual performances on this metric for clinical decision making and treatment planning purposes.

Funding:

The project described was supported by a research grant from the National Institutes on Aging (5R01AG045163). 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

ADNI2. (2020). Alzheimer's Disease Neuroimaging Initiative: ADNI2 Procedures Manual. Retrieved May 21 from https://adni.loni.usc.edu/wp-content/uploads/2008/07/adni2-procedures-manual.pdf

Benedict R (1997). Brief Visuospatial Memory Test-Revised. Psychological Assessment Resources, Inc.

Bland JM, & Altman DG (2011, Mar 11). Correlation in restricted ranges of data. BMJ, 342, d556. 10.1136/bmj.d556 [PubMed: 21398359]

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

Cherner M, Suarez P, Lazzaretto D, Fortuny LA, Mindt MR, Dawes S, Marcotte T, Grant I, Heaton R, & group H. (2007, Mar). Demographically corrected norms for the Brief Visuospatial Memory Test-revised and Hopkins Verbal Learning Test-revised in monolingual Spanish speakers from the U.S.-Mexico border region. Arch Clin Neuropsychol, 22(3), 343–353. 10.1016/j.acn.2007.01.009 [PubMed: 17293078]

Duff K (2016). Demographically corrected normative data for the Hopkins Verbal Learning Test-Revised and Brief Visuospatial Memory Test-Revised in an elderly sample. Appl Neuropsychol Adult, 23(3), 179–185. 10.1080/23279095.2015.1030019 [PubMed: 26496163]

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

Duff K, Patton D, Schoenberg MR, Mold J, Scott JG, & Adams RL (2003, Aug). Age- and education-corrected independent normative data for the RBANS in a community dwelling elderly sample. Clin Neuropsychol, 17(3), 351–366. 10.1076/clin.17.3.351.18082 [PubMed: 14704885]

Duff K, & Ramezani A (2015, Nov). Regression-Based Normative Formulae for the Repeatable Battery for the Assessment of Neuropsychological Status for Older Adults. Arch Clin Neuropsychol, 30(7), 600–604. 10.1093/arclin/acv052 [PubMed: 26289055]

Duff K, Schoenberg MR, Mold JW, Scott JG, & Adams RL (2011, Apr). Gender differences on the Repeatable Battery for the Assessment of Neuropsychological Status subtests in older adults: baseline and retest data. J Clin Exp Neuropsychol, 33(4), 448–455. 10.1080/13803395.2010.533156 [PubMed: 21154078]

Folstein MF, Folstein SE, & McHugh PR (1975, Nov). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res, 12(3), 189–198. 10.1016/0022-3956(75)90026-6 [PubMed: 1202204]

Gifford KA, Phillips JS, Samuels LR, Lane EM, Bell SP, Liu D, Hohman TJ, Romano RR 3rd, Fritzsche LR, Lu Z, Jefferson AL, & Alzheimer's Disease Neuroimaging I. (2015, Jul). Associations between Verbal Learning Slope and Neuroimaging Markers across the Cognitive Aging Spectrum. J Int Neuropsychol Soc, 21(6), 455–467. 10.1017/S1355617715000430 [PubMed: 26219209]

Gogos A, Joshua N, & Rossell SL (2010, Mar). Use of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) to investigate group and gender differences in schizophrenia and bipolar disorder. Aust N Z J Psychiatry, 44(3), 220–229. 10.3109/00048670903446882 [PubMed: 20050718]

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

Hammers DB, Duff K, & Spencer RJ (2021, Apr 26). Demographically-corrected normative data for the HVLT-R, BVMT-R, and Aggregated Learning Ratio values in a sample of older adults. J Clin Exp Neuropsychol, 1–11. 10.1080/13803395.2021.1917523 [PubMed: 33302782]

Hammers DB, Gradwohl BD, Kucera A, Abildskov T, Wilde EA, & Spencer RJ (In Press). Preliminary validation of a measure for learning slope for the HVLT-R and BVMT-R in older adults. Cognitive and Behavioral Neurology.

Hammers DB, Suhrie KR, Dixon A, Gradwohl BD, Archibald ZG, King JB, Spencer RJ, Duff K, & Hoffman JM (2021, May 5). Relationship between a novel learning slope metric and Alzheimer's disease biomarkers. Aging, Neuropsychology, and Cognition, 1–21. 10.1080/13825585.2021.1919984

Hammers DB, Suhrie KR, Dixon A, Gradwohl BD, Duff K, & Spencer RJ (2021, April 24). Validation of HVLT-R, BVMT-R, and RBANS learning slope scores along the Alzheimer's continuum. Arch Clin Neuropsychol. 10.1093/arclin/acab023

Harrington KD, Lim YY, Ames D, Hassenstab J, Rainey-Smith S, Robertson J, Salvado O, Masters CL, Maruff P, & Group A. R. (2017, Mar 1). Using Robust Normative Data to Investigate the Neuropsychology of Cognitive Aging. Arch Clin Neuropsychol, 32(2), 142–154. 10.1093/arclin/acw106 [PubMed: 27932344]

Ivnik R, Malec J, Smith G, Tangalos E, & Petersen R (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE token, WRAT-R reading, AMNART, STROOP, TMT, and JLO. Clin Neuropsychol, 10(3), 262–278. 10.1080/13854049608406689

Ivnik R, Malec J, Smith G, Tangalos E, Petersen R, & Kokmen E (1992). Mayo's older Americans normative studies: WAIS-R norms for ages 56 to 97. Clin Neuropsychol, 6, 1–30.

Lemos R, Duro D, Simoes MR, & Santana I (2014, Nov). The free and cued selective reminding test distinguishes frontotemporal dementia from Alzheimer's disease. Arch Clin Neuropsychol, 29(7), 670–679. 10.1093/arclin/acu031 [PubMed: 25062746]

Lezak MD, Howieson DB, Bigler ED, & Tranel D (2012). Neuropsycholgoical Assesssment (5th ed.). Oxford University Press.

Mast BT, & Allaire JC (2006, May). Verbal learning and everyday functioning in dementia: an application of latent variable growth curve modeling. J Gerontol B Psychol Sci Soc Sci, 61(3), P167–173. 10.1093/geronb/61.3.p167 [PubMed: 16670186]

Morris JC (1993, Nov). The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology, 43(11), 2412–2414. 10.1212/wnl.43.11.2412-a

Norman MA, Moore DJ, Taylor M, Franklin D Jr., Cysique L, Ake C, Lazarretto D, Vaida F, Heaton RK, & Group H. (2011, Aug). Demographically corrected norms for African Americans and Caucasians on the Hopkins Verbal Learning Test-Revised, Brief Visuospatial Memory Test-Revised, Stroop Color and Word Test, and Wisconsin Card Sorting Test 64-Card Version. J Clin Exp Neuropsychol, 33(7), 793–804. 10.1080/13803395.2011.559157 [PubMed: 21547817]

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

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

Sheikh JI, & Yesavage J (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. Clinical Gerontologist, 5, 165–172.

Skodzik T, Holling H, & Pedersen A (2017, Feb). Long-Term Memory Performance in Adult ADHD. J Atten Disord, 21(4), 267–283. 10.1177/1087054713510561 [PubMed: 24232170]

Smith A (1973). Symbol Digit Modalities Test. Western Psychological Services.

Spencer RJ, Gradwohl BD, Williams TF, Kordovski VM, & Hammers DB (2020, Jul 11). Developing learning slope scores for the repeatable battery for the assessment of neuropsychological status. Appl Neuropsychol Adult, 1–7. 10.1080/23279095.2020.1791870 [PubMed: 29617165]

Suhr JA (2015). Psychological Assessment: A Problem-Solving Approach. Guilford Press.

United States Bureau of Labor Statistics. (2016). Volunteering in the United States, 2015. https://www.bls.gov/news.release/volun.nr0.htm

United States Bureau of the Census. (2015). Current Population Survey, September 2014: Volunteer Supplement Inter-university Consortium for Political and Social Research. 10.3886/ICPSR36154.v1

Wechsler D (1987). Manual for the Weschler Memory Scale - Revised. The Psychological Corporation.

Wilkinson GS (1993). WRAT-3: Wide Range Achievement Test Administration Manual. Wide Range, Inc.

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

Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, & Leirer VO (1982). Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res, 17(1), 37–49. http://www.ncbi.nlm.nih.gov/pubmed/7183759 [PubMed: 7183759]

Table 1.

Demographic, cognitive, and behavioral variables for the total sample

Variable	*Mean* (*SD*)	Range
n	200	
Age (years)	74.8 (6.7)	65 – 96
Education (years)	15.8 (2.6)	8 – 20
Sex (% female)	77.5%	
Race (% Caucasian)	99.5%	
WRAT Reading	108.2 (7.4)	81 – 126
RBANS Total Scale	109.2 (12.8)	81 – 146
RBANS Immediate Memory Index	108.1 (14.3)	57 – 152
RBANS Visuospatial/Constructional Index	105.6 (14.2)	64 – 131
RBANS Language Index	104.5 (11.1)	75 – 137
RBANS Attention Index	105.6 (14.6)	72 – 138
RBANS Delayed Memory Index	108.6 (10.3)	75 – 137
HVLT-R Total Recall	56.1 (8.6)	26 – 74
HVLT-R Delayed Recall	53.7 (8.2)	27 – 67
BVMT-R Total Recall	48.5 (10.9)	20 – 75
BVMT-R Delayed Recall	51.9 (9.7)	23 – 75
Trial Making Test, Part A	51.9 (9.6)	20 – 77
Trial Making Test, Part B	50.8 (9.7)	21 – 73
Symbol Digit Modalities Test	52.8 (8.0)	27 – 73

Note: WRAT Reading = Wide Range Achievement Test – Third and Fourth Edition Reading Subtest, RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, HVLT-R = Hopkins Verbal Learning Test – Revised, BVMT-R = Brief Visuospatial Memory Test – Revised. WRAT score and RBANS scores listed as a *Standard Score*, and HVLT-R, BVMT-R, Trail Making Tests, and Digit Symbol Modality Test scores are listed as *T Scores*. All values are *Mean* (*Standard Deviation*) unless listed otherwise.

Table 2.

RBANS Learning slope scores for the total sample

Variable	*M* (*SD*)	Range	*r with age*	*r with education*
n	200			
List Learning LR	0.704 (0.258)	0.00 – 1.00	−.27**	.06
Story Memory LR	0.710 (0.296)	−1.00 – 1.00	−.08	.17*
Aggregated RBANS LR	0.710 (0.200)	0.13 – 1.00	−.20**	.17*
List Learning Trial 1	5.00 (1.5)	2 – 10	−.39**	.24**
List Learning Trial 2	6.92 (1.6)	3 – 10	−.34**	.22**
List Learning Trial 3	7.88 (1.5)	3 – 10	−.37**	.18*
List Learning Trial 4	8.43 (1.5)	5 – 10	−.36**	.15*
Story Memory Trial 1	7.68 (2.4)	2 – 12	−.24**	.14*
Story Memory Trial 2	10.65 (1.4)	4 – 12	−.12	.20**

Note: RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, LR = Learning Ratio. Please note that LR (Mean, SD) values are listed as three decimal places for the purpose of their inclusion in later equations.

* $p < .05$,

** $p < .01$.

Table 3.

Regression equations for demographically corrected scores on the List Learning LR, Story Memory LR, and the Aggregated RBANS LR.

	F(df), p, R ^{2}	**Equation**	***SE*** $_{est}$
List Learning LR	$F(2, 197) = 7.72$, $p = .001$, $R^2 = .07$	1.45 − (*age**0.01) + (*education**0.001)	0.25
Story Memory LR	$F(2, 197) = 3.10$, $p = .04$, $R^2 = .03$	0.60 − (*age**0.002) + (*education**0.018)	0.29
Aggregated RBANS LR	$F(2, 197) = 6.19$, $p = .002$, $R^2 = .06$	0.94 − (*age**0.005) + (*education**0.011)	0.20

Note: RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, LR = Learning Ratio. Age and education are both in years. SE_{est} = Standard Error of the estimate.

Table 4.

Case Example of a 69-year-old woman with 15 years of education

	List Learning	Story Memory	Aggregated RBANS
Trial 1	3	4	7
Trial 2	6	--	--
Trial 3	7	--	--
Trial 4 (or Final Trial)	9	10	19
Total Recall *T* Score	43	37	39
Observed LR Value	0.86	0.75	0.80
Predicted LR Value	0.78	0.73	0.76
Observed - Predicted Scaled Score LR Value	0.08	0.02	0.04
SE_{est}	0.25	0.29	0.20
Age/Education Corrected LR Discrepancy Z-*score* Value	0.34	0.06	0.20
Age/Education Corrected LR *T Score* Value	54	51	52

Note: RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, LR = Learning Ratio, SE_{est} = Standard Error of the Estimate of the regression equations. Predicted LR Values are derived from the regression formula from Table 3. Age/Education Corrected LR Discrepancy *Z*-score Value = Observed - Predicted LR Value/ SE_{est}.

Table 5.

Sample demographically corrected LR value *T Score* performance distribution for a 75-year-old woman with 16 years of education

List Learning LR		Story Memory LR		Aggregated RBANS LR	
LR Value	***T Score***	**LR Value**	***T Score***	**LR Value**	***T Score***
0.00	21	0.00	25	0.00	13
0.10	25	0.10	28	0.10	18
0.20	29	0.20	32	0.20	23
0.30	33	0.30	35	0.30	28
0.40	37	0.40	38	0.40	33
0.50	41	0.50	42	0.50	38
0.60	46	0.60	45	0.60	43
0.70	49	0.70	49	0.70	48
0.80	53	0.80	52	0.80	53
0.90	57	0.90	56	0.90	58
1.00	62	1.00	59	1.00	63

Note: RBANS = Repeatable Battery for the Assessment of Neuropsychological Status, LR = Learning Ratio. LR Values reflect raw LR scores.