Serial position effects in the Logical Memory Test: Loss of primacy predicts amyloid positivity

Davide Bruno[1], **Kimberly D. Mueller**[2,3,4], **Tobey Betthauser**[4,5], **Nathaniel Chin**[4,5], **Corinne D. Engelman**[3,4,6], **Bradley Christian**[4,7,8], **Rebecca L. Koscik**[3,5], **Sterling C. Johnson**[3,4,9]

[1]School Psychology, Liverpool John Moores University, Liverpool, UK

[2]Department of Communication Sciences and Disorders, University of Wisconsin-Madison, Madison, WI, USA

[3]Wisconsin Alzheimer's Institute, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA

[4]Alzheimer's Disease Research Center, University of Wisconsin School of Medicine and Public Health, USA, Madison, WI, USA

[5]Department of Medicine, University of Wisconsin- Madison, Madison, WI, USA

[6]Department of Population Health Sciences, University of Wisconsin- Madison, Madison, WI, USA

[7]Department of Medical Physics, University of Wisconsin-Madison, Madison, WI, USA

[8]Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin-Madison, Madison, WI, USA

[9]Geriatric Research Education and Clinical Center, Wm. S. Middleton Veterans Hospital, USA, Madison WI, USA

Abstract

Background: Story recall is a frequently used neuropsychological test of episodic memory with clinical populations and for screening participants in drug trials for Alzheimer's disease. However, it is unclear at this stage which underlying mechanisms confer the test its sensitivity. In this paper, we examined serial position effects, i.e., better recall for items learned early and late on a list, in story recall, and their usefulness to predict early changes associated with neurodegenerative markers.

Methods: We analysed data from the Wisconsin Registry for Alzheimer's Prevention. First, we tested whether serial position effects were present in story recall (measured with the Wechsler Memory Scale Logical Memory task; LMT) across individuals who were classified as cognitively unimpaired – stable, cognitively unimpaired – declining, or as having mild cognitive impairment (MCI).

Corresponding Author: Davide Bruno, PhD, School of Psychology, Liverpool John Moores University, Liverpool, UK, Phone: +44 (0)151 904 6320, d.bruno@ljmu.ac.uk.

Results: Our results showed clear serial position effects for all groups, except for delayed recall among individuals with MCI, where no primacy effect was observed. Second, we tested whether loss of primacy from immediate to delayed recall was associated with amyloid burden (as measured with PiB PET) in individuals that were cognitively unimpaired at baseline. We found that more primacy loss predicted amyloid positivity, above and beyond the LMT total score.

Conclusions: This report is the first to show that loss of primacy between immediate and delayed story recall is associated with amyloid burden.

Keywords

Logical Memory Test; Story Recall; PiB PET; Serial Position; Primacy; Mild Cognitive Impairment

Introduction

Sharp episodic memory decline is characteristic of Alzheimer's disease (AD) dementia, but loss of episodic memory is also a common manifestation of non-pathological cognitive decline (e.g., 1,2). Therefore, since early detection of cognitive decline associated with AD neuropathology is fundamental for identifying individuals on a trajectory to dementia, and for screening prospective participants for clinical trials, sensitivity and specificity of neuropsychological tests of episodic memory should be enhanced, whenever possible. Among the many neuropsychological tests available for episodic memory (3), Story Recall (e.g., Logical Memory III; henceforth, LMT) (4) is a popular cognitive screening tool, as demonstrated by its use in large studies such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL). LMT involves reading two brief stories to participants, who later have to recall the content of these stories. It is a relatively inexpensive test and it is also attractive to clinical practitioners due to its purported ecological validity. However, studies examining its sensitivity to early cognitive decline for persons with MCI and dementia have yielded both positive (5) and mixed results (3, 6–8). Therefore, it is important to understand what aspects of the test are more sensitive to early detection of AD biomarkers, such as β-amyloid pathology measured by positron emission tomography (PET), and whether test scoring can be improved upon.

A common pattern in tests of human memory performance is the serial position curve: performance is typically better for stimuli learned either at the beginning (primacy) or at the end (recency) of a study list, compared to the middle (e.g., 9). In particular, primacy recall appears to be affected in individuals at risk of cognitive decline (10–12), in individuals with mild cognitive impairment (MCI) (13) and in people with AD (14). However, most research in serial position effects has examined list-learning tasks, with less attention paid to serial position effects in story recall tasks, particularly within the clinical field. A notable exception is the work of Hall and Bornstein (15), who examined serial position performance in story recall (using the LMT), comparing a population of individuals with a closed head injury (mean age = 34) and age-matched controls. Their results showed that, while the clinical group presented with poorer memory overall, serial position curves were notable in both groups: primacy recall was best when compared to the other serial positions. Similarly,

Brodsky et al. (16) found serial position effects in the recall of stories in both controls and individuals with aphasia (mean age = 63), as did Leo et al. (17) (mean age = 51), who also incorporated sung stories into their paradigm.

In summary, the existing literature appears to suggest that serial position effects are present in story recall tasks, including with clinical populations. However, when examining the evidence from individuals specifically on a trajectory to AD or with the disease, the picture becomes muddier. Johnson, Storandt and Balota (18) analysed archival LMT data from individuals categorised as non-demented (0 on the Clinical Dementia Rating, CDR, scale; 47 participants), very mildly demented/MCI (0.5 on the CDR scale; 31 participants), and mildly demented (1 on the CDR scale; 39 participants). Johnson et al. did not find strong serial position effects in any of the groups, including in young, healthy controls, but observed some recency effect in the very mildly demented individuals (mean age = 74). A partial preservation of recency recall has been reported elsewhere in similar populations (e.g., 19, 20).

In this paper, we set out to examine the serial position effect in the LMT using data from the Wisconsin Registry of Alzheimer's Prevention (WRAP). Our approach was in two stages. First, we set out to establish whether there were serial position effects in the data, and whether the serial position curve differed across severity of cognitive impairment. We conceived of this first analysis as essentially exploratory in nature and hypothesis-generating: its primary aim was to identify empirically a potential serial position index that would predict subsequent amyloid burden. Second, we tested whether this serial position index was predictive of amyloid burden in cognitively intact individuals. Amyloid burden, as measured with [C-11]Pittsburgh Compound B (PiB) PET imaging was preferred to a diagnostic classification as an endpoint since the latter is partly dependent on LMT performance (see Methods, Procedure). The second analysis focused on cognitively unimpaired individuals only as identifying cognitive measures that are sensitive to β-amyloid pathology in this specific population is of considerable value.

Methods

Participants:

Data were extracted from WRAP, an ongoing longitudinal cohort study based at the University of Wisconsin – Madison (21). At the time of analysis, there were 1564 enrolled participants in WRAP, of whom 1270 were active. Further analysis subsamples are reported below. All activities for this study were approved by the ethics committees of the authors' universities, and completed in accordance with the Helsinki Declaration. All participants provided informed consent prior to testing.

Procedure:

WRAP participants complete an entry assessment including: laboratory tests, clinical measurements, and a health history and lifestyle form to collect information on demographics, self-reported medical and psychiatric status. WRAP adopts a two-tiered consensus conference method to classify individuals' cognitive status. The first tier of

review includes applying a statistical algorithm that identifies cases where impairment may be present. The second tier includes a team review of individuals flagged by the algorithm. WRAP participants reach the second tier, consisting of a consensus case conference, if they meet one or more of the following criteria: 1) the participant is performing 1.5 SDs below the mean on factor scores or individual measures of memory, executive function, language, working memory, or attention (22, 23); 2) cognitive performance on one or more tests fell below values used in other studies as cut-points for clinical MCI diagnoses (e.g., we used LMT II: Story A + Story B delayed recall <= 16 to align with the Alzheimer's Disease Neuroimaging Intitiative's use of LMT II: story A score <=8) (24); or 3) an abnormal informant report indicating subjective cognitive or functional decline. Consensus diagnoses are determined for each visit by a research team including physicians, clinical neuropsychologists, and clinical nurse practitioners based on review of cognitive, imaging, medical history, lifestyle, subjective cognitive complaints, and informant data (25). For the purposes of this paper, three consensus diagnoses were used: cognitively unimpaired – stable (CUS), cognitively unimpaired – declining (CUD), and clinical mild cognitive impairment (MCI). Of note, CUS and CUD are consistent with the NIA-AA Research Framework (26) "cognitively unimpaired" stage, and our MCI classification is consistent with their "mild cognitive impairment" description. Participants who did not meet MCI or dementia criteria, but showed evidence of subtle declines, were assigned to the CUD group.

The neuropsychological battery, comprising commonly used clinical tests (see 27, and 28 for a description of the baseline cognitive battery), also included the Wide Range Achievement Test (WRAT) Reading Test (29), as a measure of pre-morbid IQ, and the LMT as a measure of story recall. As the latter was only taken starting from WRAP visit 2, we adopted this visit throughout as our baseline. The LMT is a story recall test comprising two stories with 25 items each (e.g., Gerard, Dancing, Fifty), belonging to different semantic and lexical categories. Each story is read out to the participant, who then is asked to recall it immediately and after a delay of 25–30 minutes. Scoring followed standardized procedures (5). While some alteration from the original item is permitted (e.g., "slid off the table" is permitted in place of "fell off the table"), other items must be remembered verbatim, such as proper names or numbers.

Genotyping.

DNA was extracted from whole blood. Samples were aliquoted on 96-well plates for determination of *APOE* genotypes at Polymorphic DNA Technologies (polymorphicdna.com, Alameda, CA). For more information, (see 30). An *APOE* risk score was calculated based on the odd ratios of the ε2/ε3/ε4 genotype, as previously reported (31).

PiB PET Imaging.

Participants underwent PET scans with PiB, acquiring scan data from 0 – 70 minutes. Amyloid burden was assessed as a global cortical average PiB distribution volume ratio (DVR) and the threshold for PiB PET positivity was set at PiB   1.19 (32).

Design and analysis.

We ran two analyses. First, was a cross-sectional comparison of LMT performance across consensus diagnoses with a 2 X 3 X 3 repeated measures analysis of co-variance (ANCOVA). The independent variables were: Delay (immediate vs. delayed testing); Serial Position (primacy vs. middle vs. recency); and consensus diagnosis (CUS vs. CUD vs. MCI). Co-variates were age and WRAT, which differed across diagnostic groups (see Table 1). The dependent variable was memory performance, measured as a proportion (see below), or as the total LMT scores. For this analysis, participants were selected that fell in any of the aforementioned diagnoses, thus excluding people with dementia and/or cognitive impairment not due to MCI. We also included only participants for whom we had complete data for all variables. The analysis was cross-sectional, using data from WRAP visit 2, which allowed us to include LMT and have the largest available sample. Overall, 653 participants were included. Table 1 provides a breakdown of demographic values across groups.

While years of education was a trend (see Table 1), we preferred controlling for the WRAT score in the analysis as WRAT is less affected by ethnic differences (33), and thus can be considered a less culturally-biased proxy of cognitive reserve compared to years of education.

Primacy was defined as the first eight items in each story, middle was defined as the next nine items, and recency was defined as the final eight items. Average scores were obtained across LMT versions A and B, and scores were then converted to proportions, dividing by eight, nine and eight, respectively. Post-hoc tests were conducted with paired-samples t-tests to disambiguate interactions.

The second analysis tested whether the serial position score chosen from the empirical results of the first analysis (see Results) predicted a binary classification of amyloid positivity at the final PiB PET scan (positive, n=49; negative, n=174). For this purpose, we selected data from participants who were cognitively unimpaired at WRAP Visit 2, and were either still unimpaired or MCI at the final assessment. Participants had to have undergone at least one PiB PET scan (we used the last measurement for the analysis) and have available WRAT and LMT data. All in all, 223 participants were included in this analysis (baseline: CUS, n=177 and CUD, n=48; final assessment: CUS, n=188, CUD, n=26, and MCI, n=9). Their age at Visit 2 ranged from 44 to 71, with a mean of 58.4 and a SD of 6.1. Time between neuropsychological assessment and PiB PET scan varied from 0 to 13 years, with an average of 8.1 years and a SD of 3.1. The analysis was a logistic binary regression. Covariates/predictors included were: WRAT score at Visit 2, sex, age at last PiB PET scan, time between neuropsychological visit and PiB PET scan, APOE risk score, and one of the following: the total LMT immediate score (Visit 2), the total LMT delayed score (Visit 2), or the serial position variable (Visit 2). As some of the LMT derived scores were correlated, we conducted three separate analyses. The total LMT immediate and delayed scores are calculated from adding all the correctly recalled items in both LMT lists at immediate and delayed recall, respectively. Analyses were carried out in SPSS 25.0 and 26.0.

Data Availability.

Data were obtained from the WRAP database. WRAP data can be requested following an application to the WRAP Science Executive Committee. More information is available at this address: http://www.wai.wisc.edu/research/wrapdatarequests.html.

Results

Cross-sectional analysis.

First, we tested the ANCOVA assumptions: the covariates were not correlated (r = 0.032); the standardised residuals were largely normally distributed according to inspection of histograms; and homogeneity of variance was achieved for all but one dependent variable, as five out of six Levene tests were not significant. While these results are not perfect, they provide reasonable support to the choice of carrying out the planned ANCOVA, especially when considering its exploratory nature.

A $2 \times 3 \times 3$ ANCOVA was run. Main effects of Delay [F(1,648)=12.707, p<0.001] and Consensus Diagnosis [F(2,648)=77.922, p<0.001] were observed, while Serial Position was not significant [F(2,1296)=1.577, p=0.207]. These main effects were qualified by significant interactions: Delay X Consensus Diagnosis [F(2,648)=11.080, p<0.001], and a three-way interaction between Delay X Serial Position X Consensus Diagnosis [F(4,1296)=2.684, p=0.030], whereas the other interactions were not significant (p's ≥ 0.115). The three-way interaction is depicted in Figure 1, and means and SDs of serial position scores are presented in Table 1. Visual inspection of the interaction suggests that clear serial position effects are present in the LMT data, with better performance at primacy and recency than middle recall. Looking across diagnoses, however, we can see a progressive flattening of the delayed curve, resulting ultimately in a lack of primacy effect in the MCI group. This effect is notable if we look at the dashed line, which becomes steeper going from CUS to CUD and then MCI, indicating a (cross-sectional) progressive increase in loss of primacy recall between immediate and delayed testing. Post-hoc tests confirm that in each diagnostic category, the difference between immediate and delayed primacy is significant (p's < 0.001), but the effect size (Cohen's *d*) varies: 0.71 for CUS, 1.03 for CUD, and 2.21 for MCI (for comparison, the analogous effect sizes for recency were 0.19, 0.29 and 0.45, respectively). To test this statistically, we also computed a change-in-primacy ratio (primacy ratio score, for simplicity), using the following formula: delayed primacy / immediate primacy (no data points were lost due to zero values). Then we ran a univariate ANOVA across Consensus Diagnosis, co-varying again age and WRAT. The main effect was significant [F(2,648)=29.550, p<0.001], as were all simple effects [CUS vs. CUD: p<0.001; CUD vs. MCI, p<0.001; CUS vs. MCI, p<0.001], confirming that the primacy ratio score decreases (CUS, 0.84, SD = 0.17; CUD, 0.74, SD = 0.23; MCI, 0.53, SD = 0.32) as the consensus diagnosis becomes more severe, indexing more loss of primacy recall.

Figure 2 reports means and standard errors of total immediate and delayed LMT scores. A 2×3 (Delay X Consensus Diagnosis) ANCOVA, controlling for age at visit 2 and WRAT, showed main effects of Delay [F(1,648)=12.760, p<0.001] and Consensus Diagnosis [F(2,648)=77.469, p<0.001]; these main effects were qualified by an interaction

[F(2,648)=10.596, p<0.001]. Like with primacy, all post-hoc comparisons across immediate and delayed tests showed significant differences (p's < 0.001), which increased as cognitive status got worse cross-sectionally: effect sizes (Cohen's *d*) were 0.51 for CUS, 0.65 for CUD and 1.09 for MCI. Notably, all effect sizes are lower for total scores compared to primacy scores, in each diagnostic category.

Binomial regression analysis.

Based on the cross-sectional analysis and visual inspection of Figure 1, we posited that the difference between LMT primacy at the immediate test and LMT primacy at the delayed test was the best LMT serial position measure for the identification of individuals at increased risk of PET-detectable brain amyloid. This hypothesis was tested next using the primacy ratio score (see above). The score was not significantly affected by final diagnosis (p=0.821; CUS, mean=0.83, SD=0.17; CUD, mean=0.82, SD=0.24; MCI, mean=0.87, SD=0.17). The logistic regression model was statistically significant, $\chi^2(6) = 58.559$, p < .001. The model explained 36% (Nagelkerke r^2) of the variance. Significant predictors of PiB PET positivity were age at final PiB assessment (Wald = 19.162, p < 0.001, ORs = 1.183), indicating the older participants were more at risk of PiB PET positivity; APOE risk score (Wald = 24.997, p < 0.001, ORs = 4.137), showing that more APOE risk increased the probability of being PiB PET positive; and finally the primacy ratio score (Wald = 3.866, p = 0.049, ORs = 0.126), meaning that more primacy loss from immediate to delayed recall corresponded to a higher risk for PiB PET positivity. In contrast, neither the immediate total LMT score (p = 0.784) nor the delayed total LMT score (p = 0.401) were significant predictors of amyloid positivity status after adjusting for other covariates. Analogously, a ratio score combining immediate total LMT and delayed total LMT in a similar fashion to the primacy ratio score failed to reach significance (p = 0.243). See Figure 3 for a comparison of mean primacy ratio scores across PiB positivity groups.

Discussion

The data from WRAP were analysed in two ways. First, we examined whether serial position effects were present in LMT data, across individuals with consensus diagnoses for CUS, CUD or MCI. We observed that clear serial position effects were present, with typical primacy and recency effects, except for people with MCI in the delayed trial. This finding is important in at least two ways. First, it confirms that, even when recalling a story, which is a coherent, semantically relevant whole, participants consistently (at least when queried immediately after story presentation) still present effects of presentation order. Second, we see that primacy recall after a delay dips in people with MCI, and that generally the loss of primacy becomes more pronounced as the diagnosis becomes more severe (see Figure 1). Following this latest observation, we then tested whether a ratio indexing loss of primacy from immediate to delayed recall was helpful to predict who, among cognitively intact (CUS and CUD) participants, was at greater risk of showing higher amyloid burden in the PET scan. Indeed, losing more primacy was associated with a greater chance of PiB positivity, when controlling, among other things, for age and consensus diagnosis. Critically, we showed that the same was not true of total LMT immediate or delayed scores, which are normally used in clinical practice and research for diagnosis and screening.

Our findings are broadly comparable to previous results from the clinical literature showing serial position effects in LMT (e.g., 15), but not specifically consistent with the results reported by Johnson et al. (18), who only observed recency effects, and no primacy effects, in individuals with a CDR score of 0.5. It should be noted, however, that the analytical approach that Johnson et al. employed is rather different from our own, including in the way that items were parsed. While, for example, we counted each of the 25 words in the story as a single item, Johnson at al. created units that sometimes consisted of more than one word (e.g., "Gerard the Giraffe" as a single unit rather than two). Moreover, the diagnostic categories in our study and in Johnson et al.'s study are not identical. Replication therefore is needed to support the present findings.

The finding that the primacy effect is weaker in MCI after a delay is consistent with a series of previous reports based on list-learning performance. Bruno et al. (10), for example, showed that individuals whose delayed primacy was lower in the Rey's Auditory Verbal Learning Test (AVLT) were at greater risk for longitudinal decline, as measured by the Mini-mental State Exam test, than those with higher delayed primacy. More recently, Talamonti et al. (12), who also examined WRAP data, provided evidence that poorer delayed primacy at baseline, also measured with AVLT, predicts conversion from CUS to CUD status, longitudinally. Cross-sectionally, other reports have highlighted that individuals with MCI present with poorer primacy recall than controls (e.g., 11). However, what remains unclear is why there is a drop in primacy performance after a delay in individuals with poorer cognition, and why this drop is predictive of further decline. Previously, we have argued that two processes ought to be acknowledged in order to understand the relationship between primacy performance and cognitive decline (34). First, which items are recalled first should be considered, as where on a list recall is initiated has a strong influence of subsequent recall (35, 36). Second, once recall begins, associative information about the temporal context experienced during learning (i.e., information about the learning order) is activated, and using this information effectively is linked to better memory (37) and reduced risk of longitudinal cognitive decline (38). Both of these points are predicated upon the ability of the individual to encode, store and retrieve information defining the temporal properties of the learned items, appealing therefore to associative memory or contextual binding of temporal information. As the issues we observe in our data appear to emerge only after a delay, we can rule out explanations based on encoding deficiencies. Similarly, although it may be a retrieval-based issue, evidence from the serial order learning literature suggests this may not be the case, since individuals with MCI have issues preserving information about the temporal order of items regardless of how this is tested (39). Hence, we suggest that the drop in primacy between immediate and delayed recall, which we find to predict PET PiB positivity, may be related with a failure to consolidate effectively the associative cues, i.e., temporal context information, which then aid the retrieval of temporally clustered items (e.g., retrieval of item 1 facilitates retrieval of item 2 as they share similar encoding contexts; e.g., 40, 41). This hypothesis requires further testing.

An obvious limitation of this study is that WRAP is a relatively young and relatively homogeneous cohort, enriched for parental history of AD. We do not know how well these findings may translate to older cohorts (that are more likely to have co-pathology and non-AD-related cognitive problems), to cohorts that present more ethnic variety, or

to non-enriched cohorts. Therefore, further testing of the primacy ratio score in LMT is recommended. Another potential limitation is the partial overlap between lexical features of the Story Recall items (e.g., nouns, verbs) and serial position. For instance, we know that recall of proper names is more predictive of subsequent amyloid burden than recall of common names (42), and that proper names tend to appear more frequently at the start of the story (primacy region). Therefore, more research should be conducted to disentangle this potential conflict. Another potential limitation is that our analyses focused strictly on those with PET PiB evidence of Aβ pathology; future analyses will repeat analyses among those who have CSF biomarkers.

To summarize, this paper presents evidence that the analysis of serial position in story recall data is a valuable tool in the researcher and clinician's arsenal, as it may provide more accurate detection of β-amyloid risk than the total LMT score. The serial position measures used in this paper are easy to extract from commonly collected LMT data, a test that is currently employed worldwide and enjoys great popularity. Therefore, use of serial position measures for screening and for clinical evaluations has high acceptability (i.e., the tools themselves are already in use), it is highly accessible (i.e., inexpensive) and may eventually prove to be accurate in prediction.

Acknowledgments

References

1. Salthouse TA (2010). Selective review of cognitive aging. Journal of the International Neuropsychological Society: JINS, 16(5), 754. [PubMed: 20673381]

2. Venneri A, Mitolo M, & De Marco M. (2016). Paradigm shift: semantic memory decline as a biomarker of preclinical Alzheimer's disease.

3. Silva D, Guerreiro M, Maroco J, Santana I, Rodrigues A, Marques JB, & de Mendonça A. (2012). Comparison of four verbal memory tests for the diagnosis and predictive value of mild cognitive impairment. Dementia and geriatric cognitive disorders extra, 2(1), 120–131. [PubMed: 22590473]

4. Wechsler D. (1999). Manual for the Wechsler abbreviated intelligence scale (WASI). San Antonio, TX: The Psychological Corporation.

5. Ritchie K, Ropacki M, Albala B, Harrison J, Kaye J, Kramer J, Randolph C, Ritchie CW. Recommended cognitive outcomes in preclinical Alzheimer's disease: Consensus statement from the European Prevention of Alzheimer's Dementia project. Alzheimer's & Dementia. 201721;13(2):186–95.

6. Storandt M, Mintun MA, Head D, Morris JC. Cognitive decline and brain volume loss as signatures of cerebral amyloid-β peptide deposition identified with Pittsburgh compound B: cognitive decline associated with Aβ deposition. Archives of neurology. 2009121;66(12):1476–81. [PubMed: 20008651]

7. E Kennedy R SSchneider L RCutter G Alzheimer's Disease Neuroimaging Initiative T. Biomarker positive and negative subjects in the ADNI cohort: clinical characterization. Current Alzheimer Research. 2012121;9(10):1135–41. [PubMed: 22963265]

8. Baek Min & Kim Hyun & Kim SangYun. (2012). Comparison between the Story Recall Test and the Word-List Learning Test in Korean patients with mild cognitive impairment and early stage of Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 34, 396–404. [PubMed: 22263656]

9. Murdock BB Jr. The serial position effect of free recall. Journal of experimental psychology. 196211;64(5):482.

10. Bruno D, Reiss PT, Petkova E, Sidtis JJ, Pomara N. Decreased recall of primacy words predicts cognitive decline. Archives of clinical neuropsychology. 201317;28(2):95–103. [PubMed: 23299182]

11. La Rue A, Hermann B, Jones JE, Johnson S, Asthana S, Sager MA. Effect of parental family history of Alzheimer's disease on serial position profiles. Alzheimer's & Dementia. 200871;4(4):285–90.

12. Talamonti D, Koscik R, Johnson S, Bruno D. Predicting Early Mild Cognitive Impairment With Free Recall: The Primacy of Primacy. Archives of Clinical Neuropsychology. 2019417.

13. Howieson DB, Mattek N, Seeyle AM, Dodge HH, Wasserman D, Zitzelberger T, Jeffrey K. Serial position effects in mild cognitive impairment. Journal of clinical and experimental neuropsychology. 201138;33(3):292–9. [PubMed: 21128149]

14. Foldi NS, Brickman AM, Schaefer LA, Knutelska ME. Distinct serial position profiles and neuropsychological measures differentiate late life depression from normal aging and Alzheimer's disease. Psychiatry Research. 2003830;120(1):71–84. [PubMed: 14500116]

15. Hall S, Bornstein RA. Serial-position effects in paragraph recall following mild closed-head injury. Perceptual and motor skills. 19916;72(3_suppl):1295–8. [PubMed: 1961678]

16. Brodsky MB, McNeil MR, Doyle PJ, Fossett TR, Timm NH, Park GH. Auditory serial position effects in story retelling for non-brain-injured participants and persons with aphasia. Journal of speech, language, and hearing research. 2003.

17. Leo V, Sihvonen AJ, Linnavalli T, Tervaniemi M, Laine M, Soinila S, Särkämö T. Cognitive and neural mechanisms underlying the mnemonic effect of songs after stroke. NeuroImage: Clinical. 201911;24:101948.

18. Johnson DK, Storandt M, Balota DA. Discourse analysis of logical memory recall in normal aging and in dementia of the Alzheimer type. Neuropsychology. 20031;17(1):82. [PubMed: 12597076]

19. Bruno D, Koscik RL, Woodard JL, Pomara N, Johnson SC. The recency ratio as predictor of early MCI. International psychogeriatrics. 201812;30(12):1883–8. [PubMed: 29667564]

20. Stella Turchetta C, Perri R, Fadda L, Caruso G, Stefania De Simone M, Caltagirone C, Carlesimo GA. Forgetting Rate on the Recency Portion of a Word List Differentiates Mild to Moderate Alzheimer's Disease from Other Forms of Dementi. Journal of Alzheimer's Disease. 201811(Preprint):1–0.

21. Johnson SC, Koscik RL, Jonaitis EM, Clark LR, Mueller KD, Berman SE, Bendlin BB, Engelman CD, Okonkwo OC, Hogan KJ, Asthana S. The Wisconsin Registry for Alzheimer's Prevention: a review of findings and current directions. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 201811;10:130–42.

22. Koscik RL, La Rue A, Jonaitis EM, Okonkwo OC, Johnson SC, Bendlin BB, Hermann BP, Sager MA. Emergence of mild cognitive impairment in late middle-aged adults in the wisconsin registry for Alzheimer's prevention. Dementia and geriatric cognitive disorders. 2014;38(1–2):16–30. [PubMed: 24556849]

23. Clark LR, Koscik RL, Nicholas CR, Okonkwo OC, Engelman CD, Bratzke LC, Hogan KJ, Mueller KD, Bendlin BB, Carlsson CM, Asthana S. Mild cognitive impairment in late middle age in the Wisconsin registry for Alzheimer's prevention study: prevalence and characteristics using robust and standard neuropsychological normative data. Archives of Clinical Neuropsychology. 20161022;31(7):675–88. [PubMed: 27193363]

24. Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. (2010). Alzheimer's disease Neuroimaging Initiative (ADNI) clinical characterization. Neurology, 74, 201–209. [PubMed: 20042704]

25. Koscik RL, Berman SE, Clark LR, Mueller KD, Okonkwo OC, Gleason CE, et al. (2016). Intraindividual cognitive variability in middle age predicts cognitive impairment 8–10 years later:

Results from the Wisconsin Registry for Alzheimer's prevention. Journal of the International Neuropsychological Society, 22, 1016–1025. [PubMed: 27903330]

26. Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's & Dementia. 201841;14(4):535–62.

27. Sager MA, Hermann B, & La Rue A. (2005). Middle-aged children of persons with Alzheimer's disease: APOE genotypes and cognitive function in the Wisconsin Registry for Alzheimer's Prevention. Journal of Geriatric Psychiatry and Neurology, 18, 245–249. [PubMed: 16306248]

28. Johnson S, Koscik RL, Jonaitis EM, Clark LR, Mueller KD, Berman SE, et al. (2018). The Wisconsin Registry for Alzheimer's prevention: A review of findings and current directions. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 10, 130–142.

29. Robertson GJ (2009). Wide-Range Achievement Test. Corsini Encyclopedia of Psychology.

30. Johnson SC, La Rue A, Hermann BP, Xu G, Koscik RL, Jonaitis EM, Bendlin BB, Hogan KJ, Roses AD, Saunders AM and Lutz MW, 2011. The effect of TOMM40 poly-T length on gray matter volume and cognition in middle-aged persons with APOEε3/ε3 genotype. Alzheimer's & Dementia, 7(4), pp.456–465.

31. Darst BF, Koscik RL, Racine AM, Oh JM, Krause RA, Carlsson CM, Zetterberg H, Blennow K, Christian BT, Bendlin BB and Okonkwo OC, 2017. Pathway-specific polygenic risk scores as predictors of amyloid-β deposition and cognitive function in a sample at increased risk for Alzheimer's disease. Journal of Alzheimer's Disease, 55(2), pp.473–484.

32. Racine AM, Adluru N, Alexander AL, Christian BT, Okonkwo OC, Oh J, Cleary CA, Birdsill A, Hillmer AT, Murali D. and Barnhart TE, 2014. Associations between white matter microstructure and amyloid burden in preclinical Alzheimer's disease: a multimodal imaging investigation. NeuroImage: Clinical, 4, pp.604–614. [PubMed: 24936411]

33. Manly JJ, Jacobs DM, Touradji P, Small SA, & Stern Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. J Int Neuropsychol Soc, 8(3), 341–348. [PubMed: 11939693]

34. Bruno D, Grothe MJ, Nierenberg J, Sidtis JJ, Teipel SJ, Pomara N. Output order and variability in free recall are linked to cognitive ability and hippocampal volume in elderly individuals. Neuropsychologia. 201618;80:126–32. [PubMed: 26593881]

35. Howard MW, Kahana MJ. Contextual variability and serial position effects in free recall. Journal of Experimental Psychology: Learning, Memory, and Cognition. 19997;25(4):923.

36. Kahana MJ. Associative retrieval processes in free recall. Memory & cognition. 199611;24(1):103–9. [PubMed: 8822162]

37. Golomb JD, Peelle JE, Addis KM, Kahana MJ, Wingfield A. Effects of adult aging on utilization of temporal and semantic associations during free and serial recall. Memory & Cognition. 200871;36(5):947–56. [PubMed: 18630201]

38. Talamonti D, Koscik R, Johnson S, & Bruno D. (2020). Temporal contiguity and ageing: The role of memory organization in cognitive decline. Journal of Neuropsychology.

39. Gillis MM, Quinn KM, Phillips PAT, & Hampstead BM (2013). Impaired retention is responsible for temporal order memory deficits in mild cognitive impairment. Acta Psychologica, 143(1), 88–95. [PubMed: 23542809]

40. Howard MW, Fotedar MS, Datey AV, & Hasselmo ME (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. Psychological Review, 112, 75–116. [PubMed: 15631589]

41. Howard MW, & Kahana MJ (1999). Contextual variability and serial position effects in free recall. Journal of Experimental Psychology: Learning, Memory, & Cognition, 25, 923–941.

42. Mueller KD, Koscik RL, Du L, Bruno D, Jonaitis EM, Koscik AZ, Christian BT, Betthauser TJ, Chin NA, Hermann BP, & Johnson SC (2020). Proper names from story recall are associated with beta-amyloid in cognitively unimpaired adults at risk for Alzheimer's disease. Cortex; a journal devoted to the study of the nervous system and behavior, 131, 137–150. Advance online publication. 10.1016/j.cortex.2020.07.008 [PubMed: 32861209]
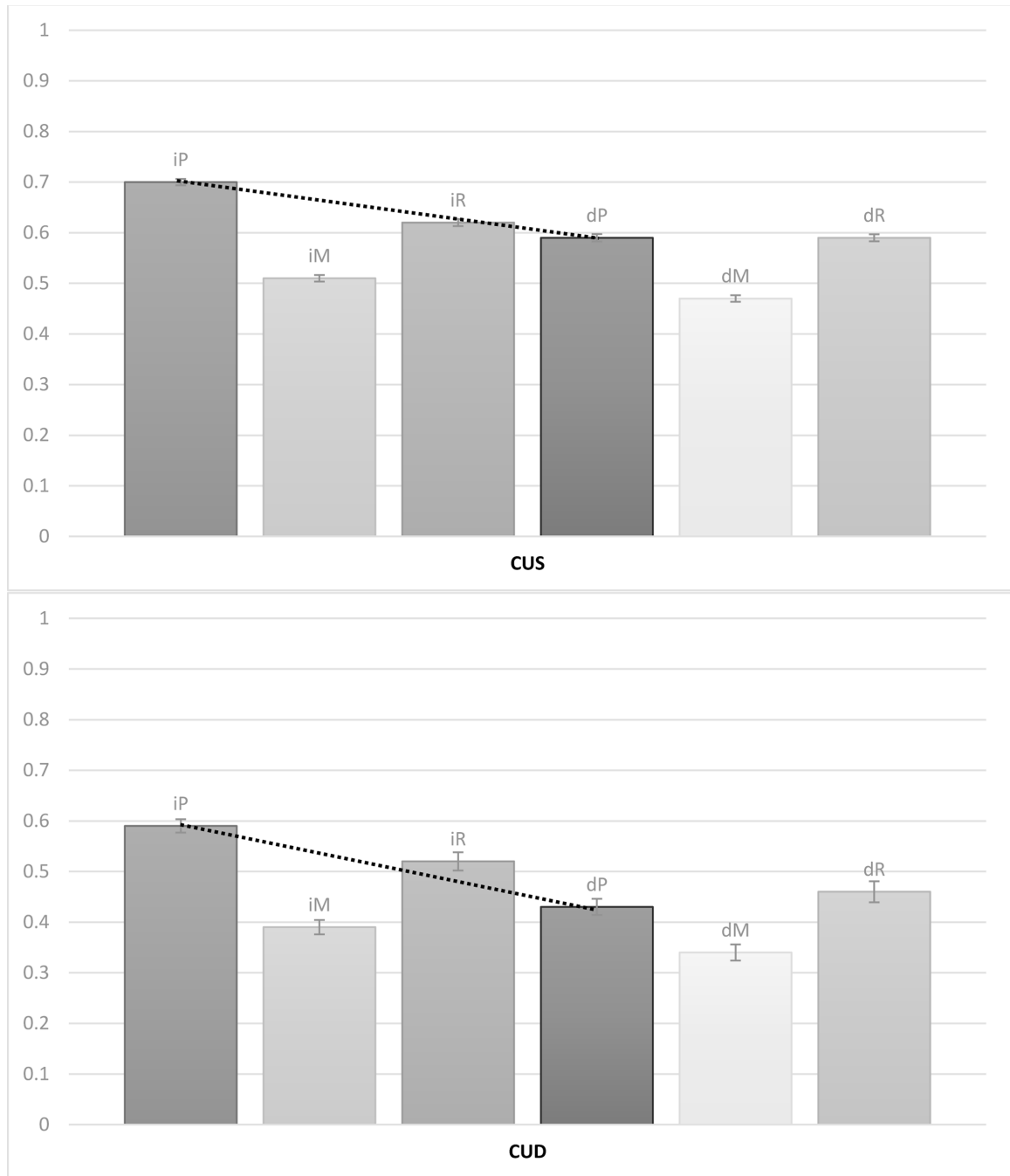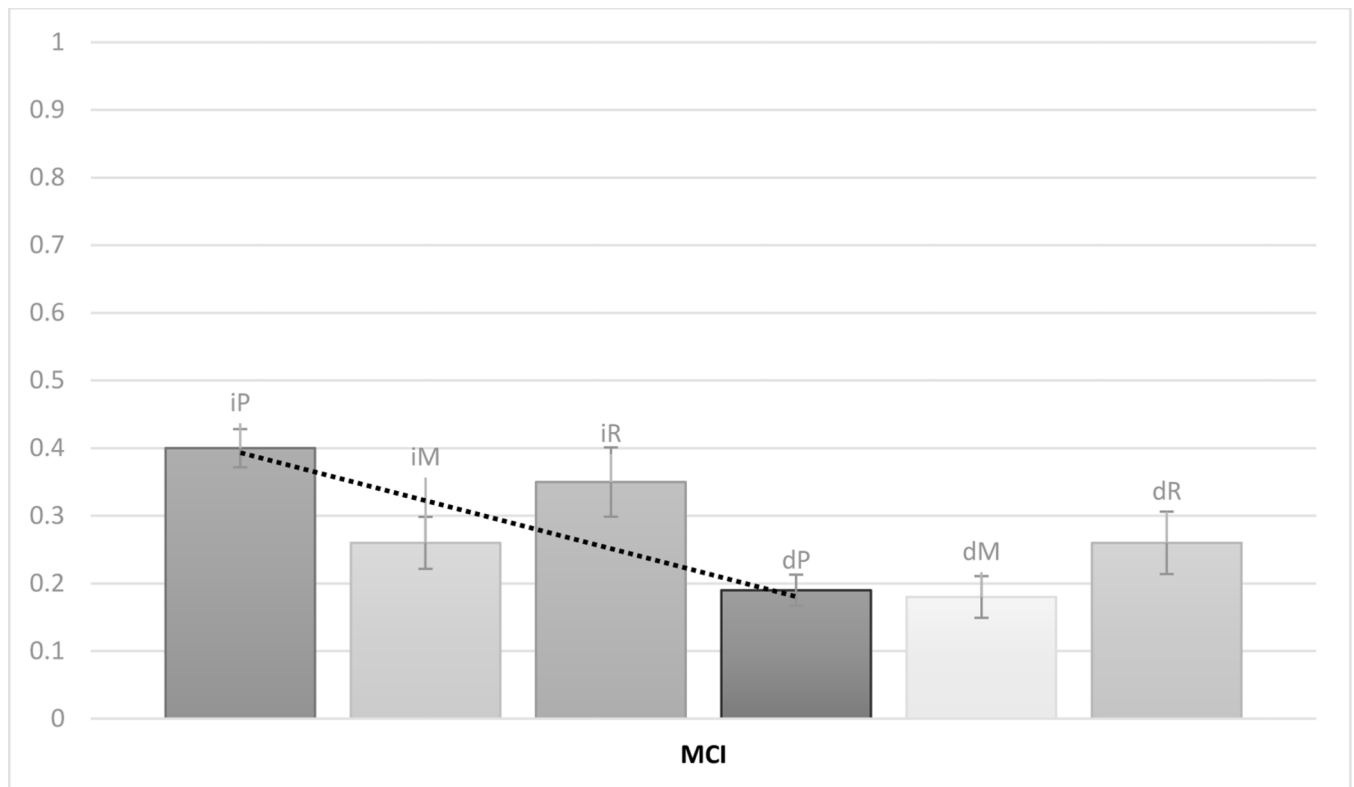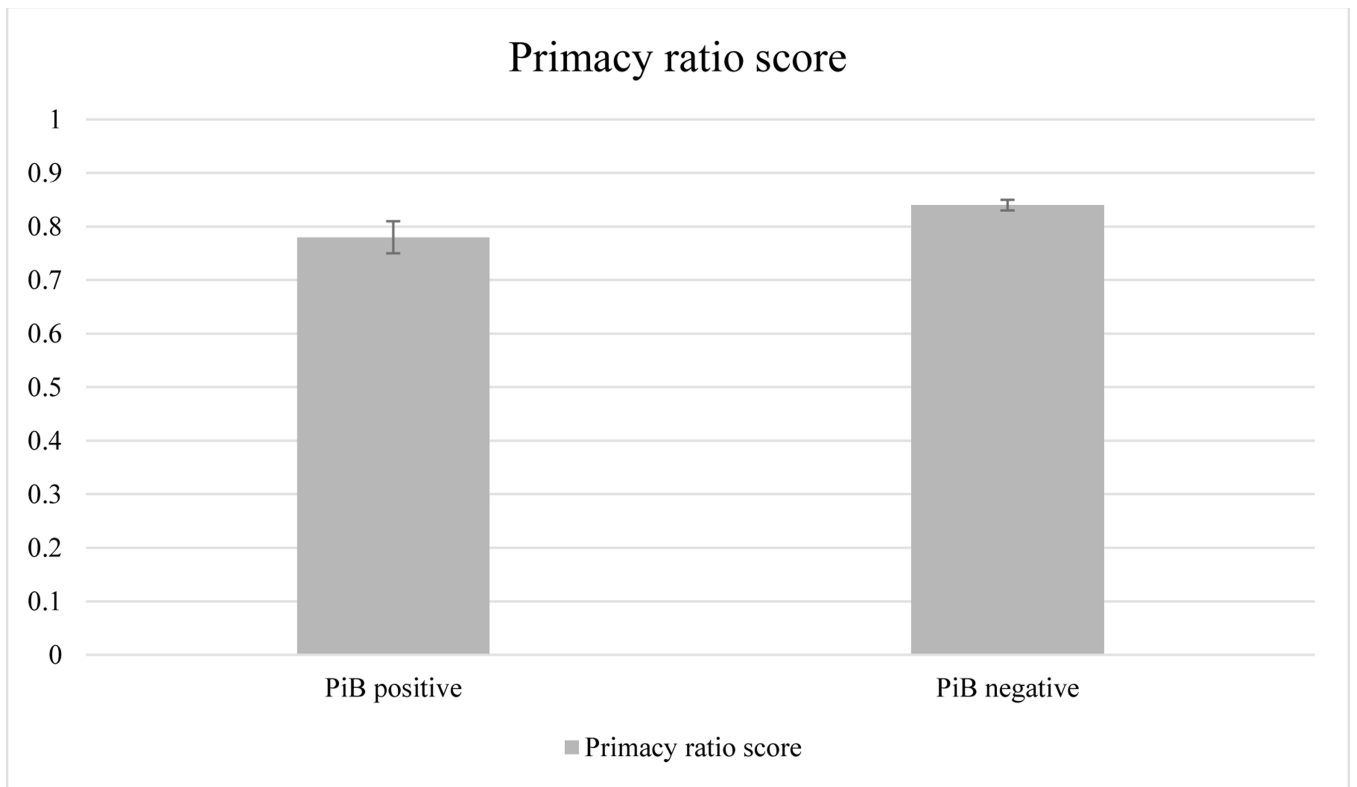
Figure 1.
Immediate primacy (iP), immediate middle (iM), immediate recency (iR), delayed primacy (dP), delayed middle (dM) and delayed recency (dR) recall proportions by consensus diagnoses: top, cognitive unimpaired – stable (CUS); middle, cognitive unimpaired – declining (CUD); and bottom, mild cognitive impairment (MCI). Numbers indicate mean score and bars indicate standard error. Dotted lines highlight the drop in primacy performance between immediate and delayed tasks within each cognitive status group.
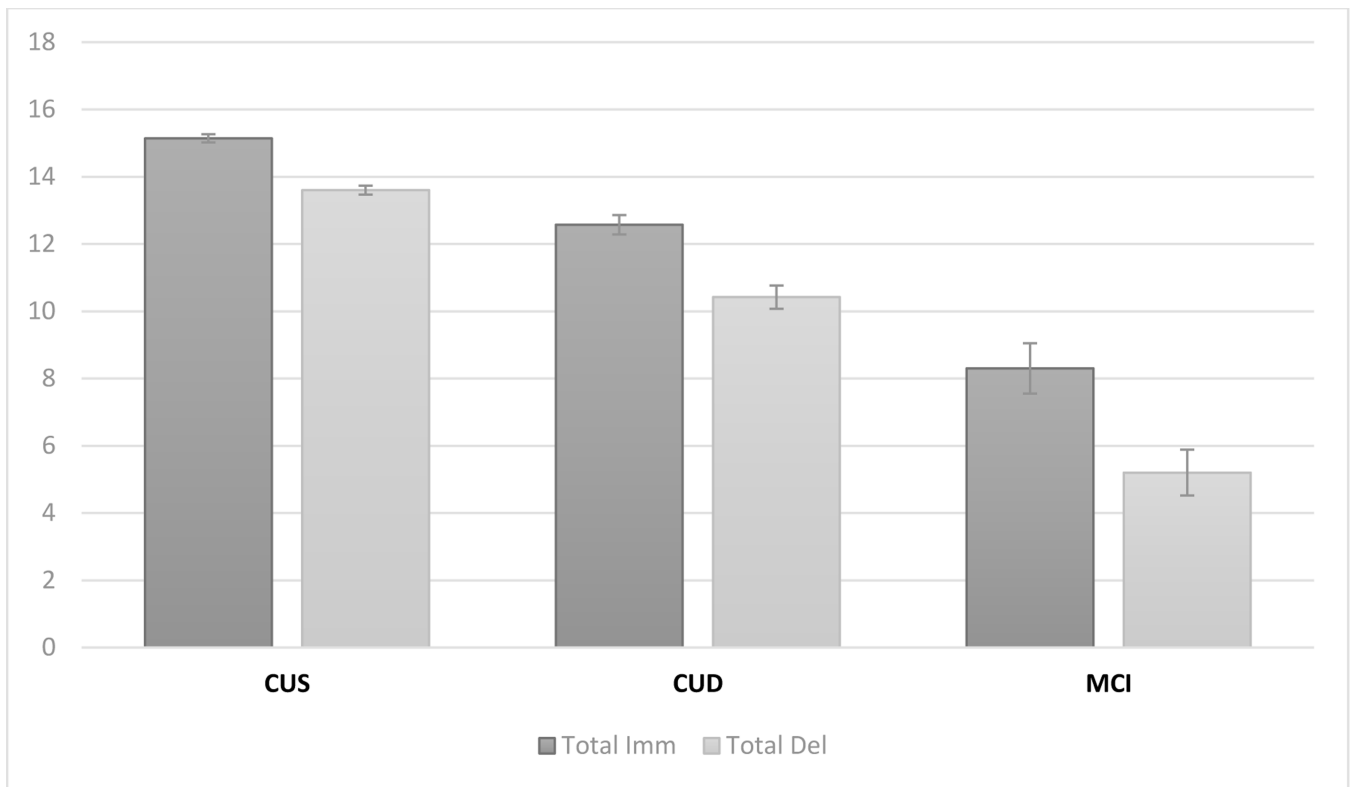
Figure 2.
Immediate (Total Imm; dark grey) and delayed (Total Del; light grey) total LMT scores by consensus diagnoses: cognitive unimpaired – stable (CUS), cognitive unimpaired – declining (CUD), and mild cognitive impairment (MCI). Numbers indicate mean score and bars indicate standard error.

Figure 3.
Mean primacy ratio score (and standard error) by PiB PET.

Table 1.

Demographic and Memory Characteristics of Study Participants by Diagnostic Group

Characteristic	Cognitively Unimpaired - Stable (N=526)	Cognitively Unimpaired - Declining (N=112)	Mild Cognitive Impairment (N=15)	p value
Age (years)	57.8 ± 6.5	60.0 ± 6.1	61.7 ± 6.7	**<0.001**
Education (years)[a]	16.3 ± 2.6	15.9 ± 2.8	14.8 ± 2.4	0.052
Females (n)	379 (72.1%)	70 (62.5%)	9 (60%)	0.091
WRAT	51.6 ± 4.3	50.4 ± 5.0	50.2 ± 3.9	**0.026**
APOE score	1.2 ± 0.7	1.2 ± 0.8	1.2 ± 0.7	0.882
				p value
Imm. Primacy	0.7 ± 0.1	0.6 ± 0.1	0.4 ± 0.1	**<0.001**
Imm. Middle	0.5 ± 0.2	0.4 ± 0.2	0.3 ± 0.1	**<0.001**
Imm. Recency	0.6 ± 0.2	0.5 ± 0.2	0.4 ± 0.2	**<0.001**
Del. Primacy	0.6 ± 0.2	0.4 ± 0.2	0.2 ± 0.1	**<0.001**
Del. Middle	0.5 ± 0.2	0.3 ± 0.2	0.2 ± 0.1	**<0.001**
Del. Recency	0.6 ± 0.2	0.5 ± 0.2	0.3 ± 0.2	**<0.001**
Imm. LMT	15.1 ± 2.8	12.6 ± 3.0	8.3 ± 3.0	**<0.001**
Del. LMT	13.6 ± 3.1	10.4 ± 3.7	5.2 ± 2.7	**<0.001**

The data are the mean ± standard deviation (SD), except for Females. Serial position scores are reported as proportions. Imm. = Immediate; Del. = Delayed; LMT = Logical Memory Test score. All tests are Univariate ANOVAs, except for Females (χ^2). Bold p values are significant (α = 0.05).