

Serial position effects reflect the integrity of distinct brain regions in Major and Minor Neurocognitive Disorder

Nancy S. Foldi^{1,*}, Kirsten I. Taylor², Andreas U. Monsch³, Marc Sollberger⁴, and Sasa L. Kivisaari⁵

*Corresponding author

¹Department of Psychology, Queens College and The Graduate Center, City University of New York, Science Building–E318, 65-30 Kissena Blvd, Queens, NY 11367 USA. tel: +01.718.997.3253, nfoldi@qc.cuny.edu

²Neuroscience, Ophthalmology, and Rare Diseases, Roche Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, 4070, Basel, Switzerland; Faculty of Psychology, University of Basel, Switzerland.

³Memory Clinic, University Center for Medicine of Aging Basel, Felix-Platter Hospital, Burgfelderstrasse 101, 4002, Basel, Switzerland.

⁴Memory Clinic, University Center for Medicine of Aging Basel, Felix Platter-Hospital, and Department of Neurology, University Hospital Basel, Burgfelderstrasse 101, 4002, Basel, Switzerland.

⁵Department of Neuroscience and Biomedical Engineering, Aalto University, School of Science, P. O. Box 12200, 00076 Aalto, Finland.

ABSTRACT

Serial Position Effects (SPE) in wordlist learning provide a rich set of metrics of cognitive functioning. In this study, we systematically mapped the neuroanatomical correlates of SPEs across delays in Major and Minor Neurocognitive disorders. Primacy, middle, and recency SPE scores of the California Verbal Learning Test at Learning, Short delay, and Long delay in healthy controls and patients with neurodegenerative disease were correlated with MRI gray matter signal intensities (voxel-based morphometry, VBM). The VBM analyses revealed distinct patterns of brain-behavioral correlations depending on both the SPE and the time-point. Unlike patients, the healthy controls' performance incrementally improved recall of primacy items from Learning to Short and to Long delay, i.e., primacy progression. Moreover, the proportion of correct primacy items recalled at long delay compared to Learning correlated with bilateral medial temporal lobe regions, which commonly bear the brunt of pathology of Alzheimer's disease. The results implicate the primacy progression score as an accessible and sensitive measure for disease detection, progression, and therapeutic response in Major and Minor Neurocognitive disorders.

Introduction

Words positioned at the beginning (i.e. primacy) and end (recency) of a word list are better recalled than words positioned in the middle of the list. These serial position effects (SPE)^{1,2} are behaviorally well-documented and putatively reflect distinct underlying cognitive processes³. Specifically, primacy item learning is thought to depend more heavily on semantic memory processes than recency item learning, which is posited to rely disproportionately on auditory rehearsal^{4,5}. This account leads to the prediction that primacy and recency recall are disproportionately affected by damage in the medial temporal lobe (MTL) and lateral temporal lobe, respectively, as these brain regions have been shown to differentially contribute to semantic memory and auditory rehearsal processes⁶⁻¹⁰. In the current voxel-based morphometry (VBM) study, we test this hypothesis in the context of Major and Minor Neurocognitive disorder. We systematically map the neuroanatomical correlates of each SPE across Learning, Short delay (SD) and Long delay (LD) recall. We predict that poorer primacy item recall is associated with decrease in gray matter volume in the MTL and suggest that this measure may serve as a sensitive marker of AD-related pathology.

Behavioral assessments of SPE at learning are well studied in healthy adults and persons with Alzheimer's disease (AD). Healthy individuals show a classic 'U-shaped' profile with comparable numbers of primacy and recency words, but fewer from the middle position. We use the term 'J-shaped'¹¹ profile to describe fewer primacy than recency items at learning, which is a well-documented characteristic performance in patients with diagnosed with AD¹¹⁻¹⁵ as well as in healthy individuals at risk for AD¹⁶⁻¹⁸. The model's divergent patterns ascribe primacy item recall to consolidation of 'deeper' semantic elaboration processes, and recency item learning to 'shallow' online phonological rehearsal¹⁹, which is relatively preserved during early stages of most forms of dementia²⁰.

Distinct cognitive subprocesses serving primacy and recency item recall at learning lead to the hypothesis that they are subserved by differential neuroanatomical bases. Namely, primacy item processing would be expected to rely on regions associated with semantic processing, such as middle (MTG) and inferior temporal gyri (ITG)¹⁰ and medial temporal lobe (MTL) structures, which support semantic memory⁸. In contrast, recency items would likely rely at least on right STG which has been linked with phonological processing^{7,19}. This anatomical differentiation has received some support in the literature. For instance, reduced primacy item learning in patients with MCI was shown to be associated with reduced gray matter intensity in e.g., temporal pole and hippocampus¹⁵.

Delayed SPE measures have been less studied, but carry a particular significance in Major and Minor Neurocognitive disorder, given that impaired episodic memory consolidation is the hallmark of Alzheimer's disease^{21,22}. Although previous studies have found associations between primacy recall at delay and hippocampal volume^{16,23} and cognitive decline²⁴, we claim more targeted steps need to be taken to grasp the full potential of the delayed primacy recall measure. First, we suggest it is important to consider the difference between performance at initial learning and subsequent delayed recalls to tap the individual rate of forgetting or of the ability to conserve primacy items over time. Second, it is important to consider brain regions beyond the hippocampus since the neighboring areas, in particular the perirhinal cortex, have been shown to carry an important role in semantic processing and item-specific memory^{6,25-27}.

The goals of the present study were to systematically determine the functional neuroanatomy of SPE at Learning, at SD, and at LD recall in healthy individuals and patients with Major and Minor Neurocognitive disorders. We anticipated that each group would show the respective U-shaped or J-shaped behavioral profiles reflecting preserved or impaired primacy item learning^{14,28}. Over time at SD and LD, only the patient group would demonstrate a drop in recency items²⁹, whereas controls would be able to maintain the initial learned primacy items. Our first hypothesis proposed that brain-behavior correlation analyses would demonstrate a reliance of primacy item learning on MTG, ITG, and MTL^{8,10} and recency item learning on the superior temporal gyrus (STG)^{7,19,30,319}. Second, we hypothesized that the range of decay or improvement of primacy items would be associated with volume in the MTL encompassing the hippocampus, entorhinal and perirhinal cortices. These are regions key regions in the early diagnosis of AD as they are affected very early in the course of AD³².

Results

Behavioral results

Analyses revealed the expected HC over PAT superiority in performance overall ($F(1,127) = 193.4, p < .001$; partial eta squared (η^2) = .60). We found a significant main effect of List Position ($F(2,254) = 3.4, p = .04$; $\eta^2 = 0.03$) with better primacy item accuracy than either middle ($p < .001$) or recency items ($p < .001$), which in turn did not differ. There also was a significant main effect of Time ($F(2,254) = 3.4, p = .03$; $\eta^2 = .03$) with significant difference between Learning and SD ($p < .05$) and LD ($p < .005$), which were not significant from one another. While age was a significant covariate ($p = .002$) and interacted with Time ($p = .03$), it did not significantly interact with SPE. The important three-way interaction was significant ($F(4,508) = 6.6, p < .001, \eta^2 = 0.05$; (Fig. 1)). At Learning: each group showed their predicted U-shaped and J-shaped learning profiles (HC: primacy = recency; PAT: primacy < recency, $p < .001$). Patients performed worse than controls on primacy and middle items (both $p < .001$), and a trend for worse performance on recency items ($p = .08$). After interference, both groups' SPE profiles change from quadratic functions at Learning to linear functions at SD and LD, where patients recalled overall fewer items than controls at all list positions ($p < .001$). Only patients, but not controls, showed the significant drop in recency item recall from Learning to SD ($p < .001$) and LD ($p < .001$), which in turn did not differ (n.s.). Importantly, the controls' primacy item performance improved significantly over time from Learning to LD ($p < .001$), while numerically improving stepwise from learning to SD ($p = .05$, trend.) and between SD to LD (n.s.). In the patient group, primacy recall remained at a level unchanged from Learning to either SD or LD (n.s.), and showed no improvement. Recency items in the patient group showed a clear decrement from Learning to SD ($p < .001$) and from Learning LD ($p < .001$), while there was no difference between the delays (SD – LD, n.s.). In contrast, the control group maintained recency item accuracy from learning to SD and to LD (n.s.).

Voxel-Based Morphometry Results

Learning

Primacy: Poor primacy item performance was associated with reduced GM volume in two clusters (Fig. 2a). One was centered in the right angular gyrus and extended to the right middle occipital gyrus. A second cluster was centered in the left MTG and included the left temporal pole, MTG/STG, hippocampus and parahippocampal gyrus. *Middle*: Poor middle item learning was associated with reduced volume in the bilateral cerebellum (see Table 1). *Recency*: Poor recency item learning was associated with reduced volume in two clusters, one centered in the left hippocampus extending to the parahippocampal cortex and amygdala. The second cluster was centered in the right superior temporal pole extending into the right parahippocampal cortex, MTG and STG (see Fig. 3). The VBM results are summarized in Table 1.

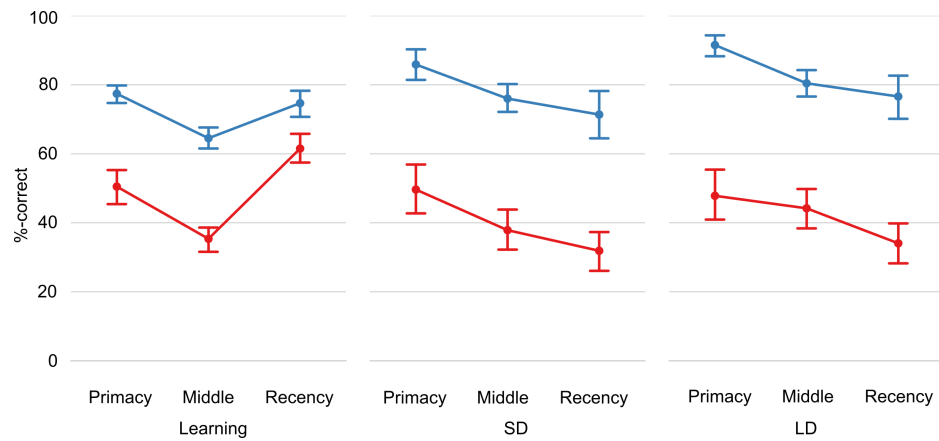


Figure 1. Performance accuracy for primacy, middle and recency serial list positions at Learning, Short Delay (SD) and Long Delay (LD) on the California Verbal Learning Test. Plots indicate mean percent correct for Healthy Controls (blue lines) and Patients (red lines). Error bars indicate 95%-confidence intervals computed using bootstrapping.

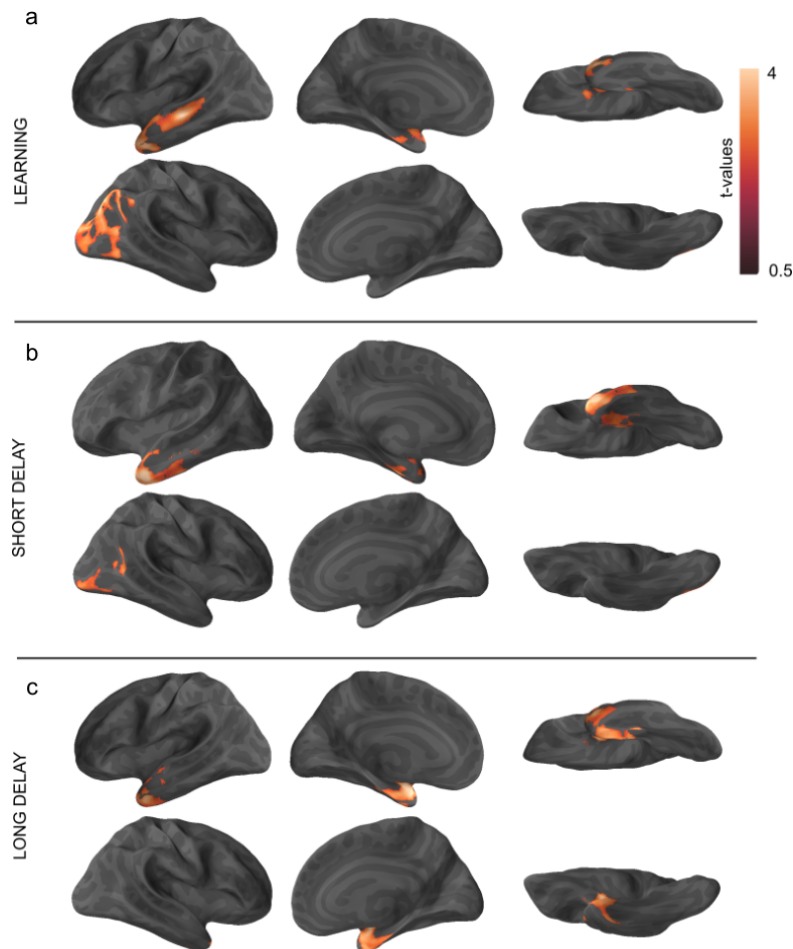


Figure 2. Brain regions where reduced signal intensity significantly correlated with poor accuracy of primacy at a) Learning, b) SD and c) LD. The t -values of voxels in significant clusters are project onto the cortical surface.

Short Delay

Primacy: Poor primacy item recall at SD was associated with reduced volumes in three clusters. One cluster was centered in the left ITG extending into the MTG and temporal pole area. A second cluster encompassed the left MTL region including the

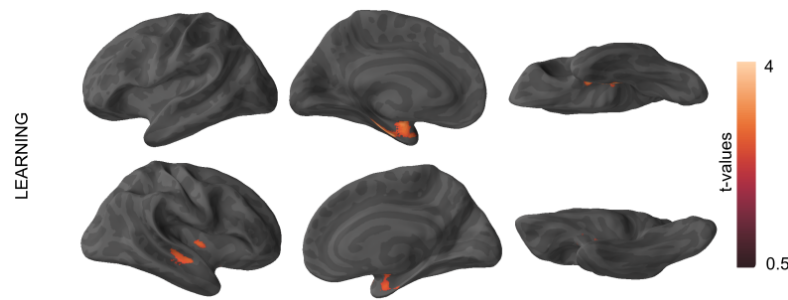


Figure 3. Brain regions where reduced signal intensity significantly correlated with poor accuracy of recency item learning. The t -values of voxels in significant clusters are projected onto the cortical surface.

hippocampus and parahippocampal gyrus. A third cluster was centered in the right occipital inferior gyrus and extended in the right MTG and angular gyrus (Fig. 2b, Table 1). *Middle*: Poor middle item recall at SD was associated with reduced volume in the bilateral cerebellum (Table 1). *Recency*: Poor recency item performance was associated with reduced bilateral inferior cerebellar volume

Long Delay

Primacy: Poor primacy item recall at LD significantly correlated with two clusters encompassing the bilateral MTL. A left-sided cluster included large portions of the amygdala, hippocampus, entorhinal cortex, and parts of the perirhinal cortex and anterior parahippocampal cortex. This cluster also extended into the fusiform gyrus, ITG, MTG and temporopolar regions. A right-sided cluster peaked in the hippocampus and extended into the amygdala, large portions of the entorhinal and perirhinal cortices and parts of the temporal pole (Fig. 2c, Table 1). *Middle*: There were no significant clusters associated with LD recall by middle items. *Recency*: There were no significant clusters associated with LD recall by recency items.

Investigation of SPE change over time: Learning to Long Delay

In the next analysis, we examined whether the change in performance of primacy, middle or recency list positions from Learning to LD was associated with gray matter signal intensity (Table 1). To address this aim, we first calculated three SPE difference scores of primacy, middle, and recency positions for each participant (%-correct at LD – %-correct at Learning, Fig. 4a). We then correlated the distribution of the difference scores with gray matter volume in three separate VBM analyses. As expected, only the difference score for primacy was significantly associated with bilateral MTL integrity indicating that poorer recall of primacy items at LD relative to Learning was associated with less volume in the MTL (Fig. 4c), Table 1. This analysis revealed no significant clusters for either the recency or middle list position across Learning to LD.

Discussion

The present study systematically studied and found support for distinct neuroanatomical correlates for SPE at Learning, SD, and LD. We found that performance with primacy items at Learning was associated with left MTL and temporal pole volume, whereas recency item Learning was associated with bilateral MTL and right STG volume. Interestingly, only primacy item recall remained significantly associated with gray matter signal intensities in the lateral and medial temporal regions over time^{16,23}. Our findings further showed that healthy controls, unlike patients, improved their item recall from Learning to SD to LD. We coin this ability to improve with these items over time ‘primacy progression’. Furthermore, reduced primacy progression was associated with smaller volumes in the bilateral MTL regions. While individual SPE measures at Learning, SD, and LD were associated with distinct neuroanatomical coordinates, we newly demonstrate that the primacy progression measure may be a particularly sensitive cognitive-biomarker differentiating controls from of incipient AD-related pathology.

Early behavioral research documented that primacy item learning differentiated healthy individuals’ U-shaped from J-shaped learning curves in patients with AD^{11,13,14,24,33}. The success of primacy item recall at learning reflects an ability to encode initial list items facilitated by binding of deeper semantic processing. In contrast, the J-shaped pattern found in MCI and AD characterizes the primacy impairment in semantic memory with preserved processing of recency items, reflecting reliance on the phonological rehearsal processes^{3,34}. Our corresponding VBM results demonstrated that poorer learning of primacy items was associated with reduced volume in the left ITG, MTG, temporal pole, MTL, and right inferior parietal areas. In previous studies, the temporal areas have been associated with semantic processing and memory^{8,10}. These findings add to the known temporal gyrus and hippocampal involvement associated with impaired primacy item recall found in AD¹⁵, amnesic MCI³⁵, left hippocampal surgical excision³⁶ and impaired hippocampal development³⁷. The findings are also consistent with an functional MRI study demonstrating heightened BOLD activity in left MTL for primacy items probes³⁸. In comparison,

	cluster-level (FDR)		peak-level (uncorrected)		coordinates			
	p	k	p	t	x	y	z	AAL
(a) Learning								
Primacy	.0001	6008	.0001	4.2	51	-64	37	Angular R
	.0001	5348	.0001	4.0	-48	-27	-8	Temporal Mid L
Middle	.0001	25220	.0001	5.0	28	-40	-51	Cerebellum R
Recency	.044	2008	.0001	4.1	-20	-10	-18	Hippocampus L
	.044	1792	.0001	3.5	46	10	-18	Temporal Pole Sup R
(b) Short delay								
Primacy	.002	3928	.0001	4.4	-48	2	-35	Temporal Inf L
	.034	1986	.0001	3.7	52	-78	-5	Occipital Inf R
	.045	1685	.0001	3.6	-14	-1	-27	Parahippocampal L
Middle	.001	3922	.0001	3.5	-6	-69	-44	Cerebellum 8 L
Recency	.005	3201	.0001	3.8	36	-45	-54	Cerebellum 8 R
	.036	2209	.0001	3.4	-30	-42	-53	Cerebellum 8 L
(c) Long delay								
Primacy	.0001	5471	.0001	4.4	-30	-4	-20	Hippocampus L
	.001	4373	.0001	4.1	27	-7	-18	Hippocampus R
Middle	ns.							
Recency	ns.							
(d) Long delay vs. Learning								
Primacy	.0001	5597	.0001	4.34	30	-10	-36	Fusiform R
	.003	3021	.0001	4.2	-33	-1	-26	Parahippocampal L
Middle	ns.							
Recency	ns.							

Table 1. VBM result summary. Overview of the VBM results demonstrating the cluster- and peak-level statistics for the three effects of interest (primacy, middle and recency) in the different VBM analyses (a) Learning, (b) Short Delay and (c) Long Delay and (d) the difference scores (Long Delay— Learning).

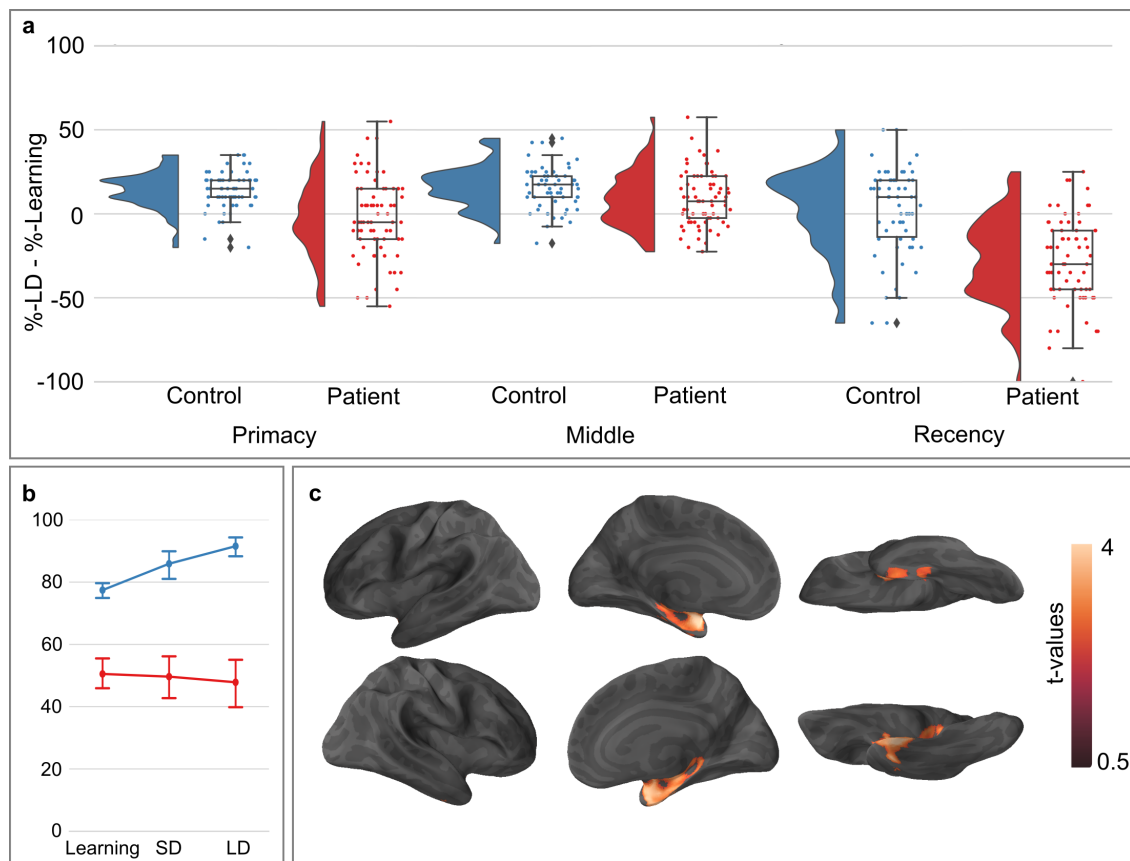


Figure 4. Primacy recall over time from Learning to LD: Primacy progression. (a) The raincloud plot reflects the difference scores for each serial list position, separately for healthy controls and patients. The half-violin plots on the left-hand side indicate the kernel density of the distribution of difference scores. The overlays of the jittered difference scores as a scatter plot and a boxplot are shown on the right-hand side. The box indicates the inter-quartile range (IQR), the band inside the box indicates the median, the whiskers indicate $1.5 \times \text{IQR}$. Outliers i.e., values $> 1.5 \times \text{IQR}$ are shown as diamonds. (b) The accuracy of primacy items at each time point (Learning, SD, and LD) for healthy controls (blue line) and patients (red line). Note that the primacy items in the control group significantly improve across three timepoints, i.e., ‘primacy progression’. In contrast, performance on primacy items in the patient group remains low and stable over time. (c) The correlation between gray matter volume and proportion of items recalled at long delay relative to those learned. This figure shows the regions where reduced volume is associated with poorer performance of delayed recalled relative to initial learning.

recency item learning was associated with reduced volume in the right STG as well as bilateral MTL. These findings reflect the importance of regions underlying auditory phonological rehearsal processes^{7,9,30,31} necessary for recency item recall³.

Two findings differentiate the neuroanatomical correlates of primacy items of SPE over time. First at SD, primacy item recall was associated with reduced left anterolateral temporal (ITG, MTG) as well as volume in the temporo-polar and right posterior temporo-occipital regions. We did not observe an association between the volume of these structures and the performance with either middle or recency items. Second, primacy item performance at LD was the only SPE measure to significantly correlate with integrity of cerebral regions. Specifically, poorer primacy item recall at LD was associated with reduced volumes of bilateral MTL, left anterolateral temporal lobe and right temporal pole. The lack of significant MTL correlation of recency items at SD or LD may reflect a floor effect given the dramatic drop in recency item performance in the patient group (see also Fig. 4a). Taken together, these findings suggest that damage to the regions, which in previous research have been shown support semantic memory processes^{8,10,39}, disproportionately hinder primacy item recall both at SD and LD as compared to other list positions.

The poor delayed recall of recency items in the patient group supports the dual process account of instability of these items owing to the shallow, phonological rehearsal processes. While patients’ and controls’ recency item performance was comparable at learning, only patients’ recency item performance dropped dramatically by roughly half at SD with no improvement at LD. In contrast, controls’ recency item performance remained stable over time. We also showed that poorer recency item

learning performance is associated with decreased bilateral MTL volume, consistent with a FDG-PET study⁴⁰, which showed that decreased metabolic activity in MTL is associated with recency item learning. Taken together, these findings provide further evidence that while phonologically processed items may be rehearsed at learning and even SD, these shallow linguistic representations are transient and difficult for patients to consolidate (see also⁴¹).

Lastly, the current study also newly demonstrates that controls, but not patients, demonstrated ‘primacy progression’, which signifies that they were able to improve item recall from Learning to LD (see Fig. 1, and Fig. 2b). We propose that this measure is a critical third serial position marker, to supplement the U- versus J-shape and the recency drop. The primacy progression denotes that at Learning, items may not have been reported but were semantically encoded and hence available on later delay. This is consistent with reports of Sederberg et al.,⁴² and Schallmo et al.,⁴³ who demonstrate that items activated at the time of encoding are those that are later reported. This is in stark contrast to patients with AD, whose poor initial primacy item recall at Learning represents not only the known semantic deficit⁴⁴ and impaired consolidation⁴⁵, but also the compounded inability to associate semantic information with the relevant ordinal position⁴⁶. Our structural brain analyses highlight the sensitivity of primacy progression and its association with integrity of the bilateral medial temporal lobes and temporal poles consistent with the semantic and memory associations. Importantly, the remaining two contrast measures for middle and recency (LD – Learning) failed to show any significant correlations with gray matter volume.

We suggest there is a two-fold explanation for the sensitivity of primacy progression to MTL pathology in Major and Minor Neurocognitive disorder. First, this measure taps both semantic and memory processes supported by parahippocampal regions^{6,8,27} and second, it also reflects consolidation processes which rely heavily on the hippocampus and entorhinal cortex⁸. The neurofibrillary pathology affects these regions early in the course of Alzheimer’s disease³² and neurofibrillary pathology associated with cognitive decline⁴⁷. Indeed, semantic memory impairments putatively reflecting pathology in these regions have previously been shown to be apparent well-before a clinical diagnosis of neurodegenerative disease⁴⁸. Thus, future studies should consider the sensitivity of the primacy progression score as an early cognitive-biomarker of Alzheimer’s disease.

Word list learning is one of the most frequently used neuropsychological instruments to measure impaired episodic memory and to detect early MTL pathology in brain disease, and AD in particular. As the CVLT list-learning measure was designed to include semantic categorization, we were able to capture focal pathology reflecting the combined impoverished semantic system and verbal episodic encoding memory networks.

We note that several different SPE scoring methods exist (see comparative review,¹⁴), but the present study offers continued empirical validation of the sensitivity of our current regional scoring method^{11,24,36,49} to regional atrophy of key neuroanatomic systems affected in neurodegenerative disease.

Limitations of the current study and future directions are worth noting. While this cross-sectional study provides a first step of cognitive-volumetric relationships of SPE profiles, future studies can delineate the longitudinal course of these measures. Future samples may also be selected to include more diverse populations, and stratify by sex. The current study addresses the SPE relationship to volume, but not to other known biomarkers associated with Alzheimer’s disease and other dementias.

The current study provides a first documentation of the functional and the neurostructural associations of primacy, middle, and recency SPE across time from initial learning to short delay followed by long delay recall. Importantly, we show that the primacy progression measure (primacy items recalled at Delay compared to primacy items correctly produced at Learning) represents a meaningful and highly sensitive cognitive-biomarker reflecting MTL pathology, which is known to be the hallmark of the development of Alzheimer’s disease. These measures provide promise for disease detection, progression, and therapeutic response.

Methods

Participants

Data of 236 individuals from longitudinal studies⁵⁰ at the Basel Memory Clinic were retrieved. All individuals were over 50-years old, scored $\geq 20/30$ on the Mini Mental Status Examination (MMSE)⁵¹ had reliable informants, and no psychiatric, neurological (other than meeting criteria for conditions below), or systemic disorders affecting cognition. Those individuals who had completed the California Verbal Learning Test (CVLT)⁵² (German version) and who underwent anatomic MRI scanning within two months (mean \pm SD) = 25 ± 14 days) were selected. Healthy Control participants (HC: $n=62$; mean age = 70.2 ± 8.3 yrs.; education = 13.2 ± 2.6 yrs.; 66 % male; mean MMSE = 29.2 ± 0.9 , range 27-30) were included if they had no neurological diagnosis and if no cognitive score on the Consortium to Establish a Registry for Alzheimer’s Disease (CERAD⁵³) Neuropsychological Assessment Battery fell below demographically-adjusted z-scores of -1.96 at baseline and at two-year follow-up, resulting in a ‘robust’ normative sample⁵⁴. Patient participants (PAT: $n=69$; mean age = 62.2 ± 9.4 yrs.; education = 13.5 ± 3.3 yrs.; 53 % male; MMSE = 28.0 ± 2.1 , range 18-30) met criteria for ICD-10 criteria for Major or Minor neurocognitive disorder. Demographics of age ($t(129) = 0.6$, ns), education ($t(129) = -0.6$, ns.) and sex ($\chi^2(1) = 2.1$, ns.) were not significantly different between groups. To increase whole-brain variability of voxel-based estimates of

gray matter (GM) volume, and thus validity of the brain-behavior correlations⁵⁵, we included patients meeting criteria using classification of Major or Minor Neurocognitive disorder. The sample included Major Neurocognitive Disorder ($n = 18$)⁵⁶; Minor Neurocognitive disorder ($n = 46$; including 44 amnesic MCI and 2 non-amnesic MCI)⁵⁷; Lewy-Body Dementia ($n = 3$)⁵⁸; Posterior Cortical Disorder ($n = 1$)⁵⁹; and frontotemporal lobar degeneration ($n = 1$)⁶⁰.

The Institutional Ethics Committees of Both Basel Cantons following principles outlined in the declaration of Helsinki and the Office of Regulatory Compliance at Queens College, City University of New York approved the study. All participants gave written informed consent prior to participation, and were reimbursed for their efforts.

Measures

Regional Serial Position Measures

The CVLT⁵² 16-item word list A was divided into primacy, middle and recency list positions (4, 8, and 4 words, respectively). Serial position scores^{11,36} were defined as items correctly recalled relative to the number presented from each list position at three time points: Learning defined as the cumulative score across Trials 1-5, SD, and LD. Note that the regional SPE scoring¹¹ differs from standard CVLT-II serial position scoring⁵².

MRI image acquisition

High-resolution T1-weighted three-dimensional magnetization-prepared rapid acquisition gradient echo images were acquired for all participants on the same 3T MRI scanner (MAGNETOM Allegra, Siemens) using the following acquisition parameters: TI = 1000 ms, TR = 2150 ms, TE = 3.49 ms, flip angle = 7°; rectangular field of view = 87.5 % (280 x 245 mm²), acquisition matrix = 256 x 224 mm, voxel size = 1.1 mm isotropic.

Design and analyses

Behavioral analyses

Age and sex were used as a covariates in the factorial analyses to be more conservative and parallel the VBM analyses. All Accuracy scores were submitted to a mixed model analysis of covariance with group (HC, PAT) as the between-subject variable, and list position (primacy, middle, recency) and time (Learning, SD, LD) as within-subjects variables. Tukey post-hoc test significance was set at $\alpha = .05$. The behavioral results were plotted using Python 3 (<https://www.python.org/>) together with Seaborn (<https://github.com/mwaskom/seaborn>) and the ptitprince libraries (<https://github.com/pog87/PtitPrince>,⁶¹).

Imaging analyses

Preprocessing and statistical analyses were conducted using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/software/>) in Matlab (v2010b, Mathworks Inc.). Anatomic images were segmented into GM volumes, which were used to create a study-specific template using the DARTEL approach⁶². Individual GM images were coregistered with the study-specific template and MNI space, modulated and smoothed with an 8 mm Gaussian kernel. Each participant's performance scores were correlated with voxel signal intensities across whole brain GM volumes using multiple regression analyses based on the General Linear Model^{55,63}. Three separate models tested the neuroanatomical correlates of primacy, middle and recency item performance at (1) Learning, (2) SD and (3) LD. Each model included performance on all three list positions. All analyses included age, sex and total GM volume as covariates. The voxel-based correlations therefore reflect the unique relationship between performance of a single list position and regional GM volume. We report contrasts theoretically relevant to our hypotheses, namely significant positive correlations for primacy, middle and recency item performance and regional brain volumes at each time point. All analyses included total GM (to control for combined effects of brain size and overall atrophy), age and sex as covariates. For each model, the t-statistic tested for regional effects, and associated p-values were corrected for search volume with the Gaussian random field theory. Statistical parametric maps were thresholded at FDR-corrected $p < .01$ at cluster-level and uncorrected at the voxel level. The anatomical areas are described using the AAL atlas (⁶⁴ and MTL subregions⁶⁵). The t-values of significant clusters are projected onto the cortical surface using Freesurfer (<https://surfer.nmr.mgh.harvard.edu/>) and PySurfer (<https://github.com/nipy/PySurfer/>) library.

References

1. Ebbinghaus, H. *Memory*. New York: Columbia University (Dover, 1885).
2. Murdock, B. B. The serial position effect of free recall. *J. experimental psychology* **64**, 482, DOI: [10.1037/h0045106](https://doi.org/10.1037/h0045106) (1962).
3. Atkinson, R. C. & Deacon, R. M. Human memory: A proposed system and its control processes. In Spence, K. W. & Spence, J. T. (eds.) *The psychology of learning and motivation: Advances in research and theories Academic Press*, vol. 2, 89–195, DOI: [10.1016/S0079-7421\(08\)60422-3](https://doi.org/10.1016/S0079-7421(08)60422-3) (Academic Press, New York, 1968).

4. Glanzer, M. & Cunitz, A. R. Two storage mechanisms in free recall. *J. verbal learning verbal behavior* **5**, 351–360, DOI: [10.1016/S0022-5371\(66\)80044-0](https://doi.org/10.1016/S0022-5371(66)80044-0) (1966).
5. Murdock, B. B. Serial order effects in short-term memory. *J. Exp. Psychol.* **76**, 1, DOI: [10.1037/h0025694](https://doi.org/10.1037/h0025694) (1968).
6. Kivisaari, S., Tyler, L. K., Monsch, A. U. & Taylor, K. I. Medial perirhinal cortex disambiguates confusable objects. *Brain* **135**, 3757–3769, DOI: [10.1093/brain/aws277](https://doi.org/10.1093/brain/aws277) (2012).
7. Hickok, G., Buchsbaum, B., Humphries, C. & Muftuler, T. Auditory–motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* **15**, 673–682, DOI: [10.1162/jocn.2003.15.5.673](https://doi.org/10.1162/jocn.2003.15.5.673) (2003).
8. Hirni, D. I., Kivisaari, S. L., Monsch, A. U. & Taylor, K. I. Distinct neuroanatomical bases of episodic and semantic memory performance in Alzheimer’s disease. *Neuropsychologia* **51**, 930–937, DOI: [10.1016/j.neuropsychologia.2013.01.013](https://doi.org/10.1016/j.neuropsychologia.2013.01.013) (2013).
9. Osaka, M. *et al.* The neural basis of individual differences in working memory capacity: an fMRI study. *NeuroImage* **18**, 789–797, DOI: [10.1016/S1053-8119\(02\)00032-0](https://doi.org/10.1016/S1053-8119(02)00032-0) (2003).
10. Binder, J. R. & Desai, R. H. The neurobiology of semantic memory. *Trends Cogn. Sci.* **15**, 527–536, DOI: [10.1016/j.tics.2011.10.001](https://doi.org/10.1016/j.tics.2011.10.001) (2011).
11. Foldi, N. S., Brickman, A. M., Schaefer, L. A. & Knutelska, M. E. Distinct serial position profiles and neuropsychological measures differentiate late life depression from normal aging and Alzheimer’s disease. *Psychiatry Res.* **120**, 71–84, DOI: [10.1016/S0165-1781\(03\)00163-X](https://doi.org/10.1016/S0165-1781(03)00163-X) (2003).
12. Buschke, H. *et al.* Retention weighted recall improves discrimination of Alzheimer’s disease. *J. Int. Neuropsychol. Soc.* **12**, 436–440, DOI: [10.1017/S135561770606053X](https://doi.org/10.1017/S135561770606053X) (2006).
13. Cunha, C., Guerreiro, M., de Mendonça, A., Oliveira, P. E. & Santana, I. Serial position effects in Alzheimer’s disease, mild cognitive impairment, and normal aging: Predictive value for conversion to dementia. *J. Clin. Exp. Neuropsychol.* **34**, 841–852, DOI: [10.1080/13803395.2012.689814](https://doi.org/10.1080/13803395.2012.689814) (2012).
14. Howieson, D. B. *et al.* Serial position effects in mild cognitive impairment. *J. Clin. Exp. Neuropsychol.* **33**, 292–299, DOI: [10.1080/13803395.2010.516742](https://doi.org/10.1080/13803395.2010.516742) (2011).
15. Kasper, E. *et al.* Neuronal correlates of serial position performance in amnesic mild cognitive impairment. *Neuropsychology* **30**, 906–914, DOI: [10.1037/neu0000287](https://doi.org/10.1037/neu0000287) (2016).
16. Bruno, D. *et al.* A study on the specificity of the association between hippocampal volume and delayed primacy performance in cognitively intact elderly individuals. *Neuropsychologia* **69**, 1–8, DOI: [10.1016/j.neuropsychologia.2015.01.025](https://doi.org/10.1016/j.neuropsychologia.2015.01.025) (2015).
17. LaRue, A. *et al.* Effect of parental family history of Alzheimer’s disease on serial position profiles. *Alzheimer’s & Dementia* **4**, 285–290, DOI: [10.1016/j.jalz.2008.03.009](https://doi.org/10.1016/j.jalz.2008.03.009) (2008).
18. Suhr, J., Demireva, P. & Heffner, K. The relation of salivary cortisol to patterns of performance on a word list learning task in healthy older adults. *Psychoneuroendocrinology* **33**, 1293–1296, DOI: [10.1016/j.psyneuen.2008.07.007](https://doi.org/10.1016/j.psyneuen.2008.07.007) (2008).
19. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393, DOI: [10.1038/nrn2113](https://doi.org/10.1038/nrn2113) (2007).
20. Huntley, J. D. & Howard, R. J. Working memory in early Alzheimer’s disease: a neuropsychological review. *Int. J. Geriatr. Psychiatry* **25**, 121–132, DOI: [10.1002/gps.2314](https://doi.org/10.1002/gps.2314) (2010).
21. Hodges, J. R. & Patterson, K. Is semantic memory consistently impaired early in the course of Alzheimer’s disease? Neuroanatomical and diagnostic implications. *Neuropsychologia* **33**, 441–459, DOI: [10.1016/0028-3932\(94\)00127-B](https://doi.org/10.1016/0028-3932(94)00127-B) (1995).
22. Martin, A. & Fedio, P. Word production and comprehension in Alzheimer’s disease: The breakdown of semantic knowledge. *Brain language* **19**, 124–141, DOI: [10.1016/0093-934X\(83\)90059-7](https://doi.org/10.1016/0093-934X(83)90059-7) (1983).
23. Bruno, D. *et al.* The relationship between cerebrospinal fluid tau markers, hippocampal volume, and delayed primacy performance in cognitively intact elderly individuals. *Alzheimer’s & Dementia: Diagn. Assess. & Dis. Monit.* **1**, 81–86, DOI: [10.1016/j.dadm.2014.11.002](https://doi.org/10.1016/j.dadm.2014.11.002) (2015).
24. Bruno, D., Reiss, P. T., Petkova, E., Sidtis, J. J. & Pomara, N. Decreased recall of primacy words predicts cognitive decline. *Arch. Clin. Neuropsychol.* **28**, 95–103, DOI: [10.1093/arclin/acs116](https://doi.org/10.1093/arclin/acs116) (2013).
25. Barense, M. D., Gaffan, D. & Graham, K. S. The human medial temporal lobe processes online representations of complex objects. *Neuropsychologia* **45**, 2963–2974, DOI: [10.1016/j.neuropsychologia.2007.05.023](https://doi.org/10.1016/j.neuropsychologia.2007.05.023) (2007).

26. Clarke, A., Devereux, B. J., Randall, B. & Tyler, L. K. Predicting the Time Course of Individual Objects with MEG. *Cereb. Cortex* DOI: [10.1093/cercor/bhu203](https://doi.org/10.1093/cercor/bhu203) (2014).
27. Kivisaari, S. L. *et al.* Reconstructing meaning from bits of information. *Nat. communications* **10**, 927, DOI: [10.1038/s41467-019-08848-0](https://doi.org/10.1038/s41467-019-08848-0) (2019).
28. Rundus, D. Analysis of rehearsal processes in free recall. *J. experimental psychology* **89**, 63, DOI: [10.1037/h0031185](https://doi.org/10.1037/h0031185) (1971).
29. Bruno, D., Kosciak, R. L., Woodard, J. L., Pomara, N. & Johnson, S. C. The recency ratio as predictor of early MCI. *Int. psychogeriatrics* 1–6, DOI: [10.1017/S1041610218000467](https://doi.org/10.1017/S1041610218000467) (2018).
30. Arnott, S. R., Grady, C. L., Hevenor, S. J., Graham, S. & Alain, C. The functional organization of auditory working memory as revealed by fMRI. *J. cognitive neuroscience* **17**, 819–831, DOI: [10.1162/0898929053747612](https://doi.org/10.1162/0898929053747612) (2005).
31. Mesgarani, N., Cheung, C., Johnson, K. & Chang, E. F. Phonetic feature encoding in human superior temporal gyrus. *Science* **343**, 1006, DOI: [10.1126/science.1245994](https://doi.org/10.1126/science.1245994) (2014).
32. Braak, H. & Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* **82**, DOI: [10.1007/BF00308809](https://doi.org/10.1007/BF00308809) (1991).
33. Pepin, E. P. & Eslinger, P. J. Verbal memory decline in Alzheimer's disease: A multiple-processes deficit. *Neurology* **39**, 1477–1477, DOI: [10.1212/WNL.39.11.1477](https://doi.org/10.1212/WNL.39.11.1477) (1989).
34. Shallice, T. & Papagno, C. Impairments of auditory-verbal short-term memory: Do selective deficits of the input phonological buffer exist? *Cortex* DOI: [10.1016/j.cortex.2018.10.004](https://doi.org/10.1016/j.cortex.2018.10.004) (2018).
35. Brueggen, K. *et al.* The primacy effect in amnesic mild cognitive impairment: associations with hippocampal functional connectivity. *Front. aging neuroscience* **8**, 244, DOI: [10.3389/fnagi.2016.00244](https://doi.org/10.3389/fnagi.2016.00244) (2016).
36. Hermann, B. P. *et al.* The effects of human hippocampal resection on the serial position curve. *Cortex* **32**, 323–334, DOI: [10.1016/S0010-9452\(96\)80054-2](https://doi.org/10.1016/S0010-9452(96)80054-2) (1996).
37. Baddeley, A., Vargha-Khadem, F. & Mishkin, M. Preserved recognition in a case of developmental amnesia: Implications for the acquisition of semantic memory? *J. Cogn. Neurosci.* **13**, 357–369, DOI: [10.1162/08989290151137403](https://doi.org/10.1162/08989290151137403) (2001).
38. Talmi, D., Grady, C. L., Goshen-Gottstein, Y. & Moscovitch, M. Neuroimaging the serial position curve: A test of single-store versus dual-store models. *Psychol. Sci.* **16**, 716–723, DOI: [10.1111/j.1467-9280.2005.01601.x](https://doi.org/10.1111/j.1467-9280.2005.01601.x) (2005).
39. Mummery, C. J. *et al.* A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Ann Neurol* **47**, 36–45, DOI: [10.1002/1531-8249\(200001\)47:1<36::AID-ANA8>3.0.CO;2-L](https://doi.org/10.1002/1531-8249(200001)47:1<36::AID-ANA8>3.0.CO;2-L) (2000).
40. Staffaroni, A. M. *et al.* The functional neuroanatomy of verbal memory in Alzheimer's disease:[18f]-Fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) correlates of recency and recognition memory. *J. clinical experimental neuropsychology* **39**, 682–693, DOI: doi.org/10.1080/13803395.2016.1255312 (2017).
41. Wolk, D. A., Dickerson, B. C. & Initiative, A. D. N. Fractionating verbal episodic memory in Alzheimer's disease. *Neuroimage* **54**, 1530–1539, DOI: [10.1016/j.neuroimage.2010.09.005](https://doi.org/10.1016/j.neuroimage.2010.09.005) (2011).
42. Sederberg, P. B. *et al.* Oscillatory correlates of the primacy effect in episodic memory. *NeuroImage* **32**, 1422–1431, DOI: [10.1016/j.neuroimage.2006.04.223](https://doi.org/10.1016/j.neuroimage.2006.04.223) (2006).
43. Schallmo, M.-P. *et al.* A new semantic list learning task to probe functioning of the Papez circuit. *J. clinical experimental neuropsychology* **37**, 816–833, DOI: [10.1080/13803395.2015.1052732](https://doi.org/10.1080/13803395.2015.1052732) (2015).
44. Monsch, A. U. *et al.* Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Arch. neurology* **49**, 1253–1258, DOI: [10.1001/archneur.1992.00530360051017](https://doi.org/10.1001/archneur.1992.00530360051017) (1992).
45. Nadel, L. & Moscovitch, M. Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr Opin Neurobiol* **7**, 217–27, DOI: [10.1016/S0959-4388\(97\)80010-4](https://doi.org/10.1016/S0959-4388(97)80010-4) (1997).
46. Ginsburg, V., Archambeau, K., van Dijck, J.-P., Chetail, F. & Gevers, W. Coding of serial order in verbal, visual and spatial working memory. *J. Exp. Psychol. Gen.* **146**, 632, DOI: [10.1037/xge0000278](https://doi.org/10.1037/xge0000278) (2017).
47. Riley, K. P., Snowden, D. A. & Markesbery, W. R. Alzheimer's neurofibrillary pathology and the spectrum of cognitive function: Findings from the Nun Study. *Annals Neurol.* **51**, 567–577, DOI: [10.1002/ana.10161](https://doi.org/10.1002/ana.10161) (2002).
48. Hirni, D. I. *et al.* Neuropsychological markers of medial perirhinal and entorhinal cortex functioning are impaired twelve years preceding diagnosis of Alzheimer's dementia. *J. Alzheimer's Dis.* **52**, 573–580, DOI: [10.3233/JAD-150158](https://doi.org/10.3233/JAD-150158) (2016).

49. Egli, S. C. *et al.* Serial position effects are sensitive predictors of conversion from MCI to Alzheimer's disease dementia. *Alzheimer's & Dementia* **10**, S420–S424, DOI: [10.1016/j.jalz.2013.09.012](https://doi.org/10.1016/j.jalz.2013.09.012) (2014).
50. Zehnder, A. *et al.* Impact of APOE status on cognitive maintenance in healthy elderly persons. *Int. J. Geriatr. Psychiatry* **24**, 132–141, DOI: [10.1002/gps.2080](https://doi.org/10.1002/gps.2080) (2009).
51. Folstein, M. F., Folstein, S. E. & McHugh, P. R. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198, DOI: [10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6) (1975).
52. Delis, D. C. California verbal learning test-second edition. *Adult version. Manual. Psychol. Corp.* (2000).
53. Welsh, K. A. *et al.* The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology* **44**, 609, DOI: [10.1212/WNL.44.4.609](https://doi.org/10.1212/WNL.44.4.609) (1994).
54. Holtzer, R. *et al.* Robust norms for selected neuropsychological tests in older adults. *Arch. Clin. Neuropsychol.* **23**, 531–541, DOI: [10.1016/j.acn.2008.05.004](https://doi.org/10.1016/j.acn.2008.05.004) (2008).
55. Tyler, L. K., Marslen-Wilson, W. & Stamatakis, E. A. Dissociating neuro-cognitive component processes: voxel-based correlational methodology. *Neuropsychologia* **43**, 771–778, DOI: [10.1016/j.neuropsychologia.2004.07.020](https://doi.org/10.1016/j.neuropsychologia.2004.07.020) (2005).
56. McKhann, G. M. *et al.* The diagnosis of dementia 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* **7**, 263–269, DOI: [10.1016/j.jalz.2011.03.005](https://doi.org/10.1016/j.jalz.2011.03.005) (2011).
57. Albert, M. S. *et al.* 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* **7**, 270–279, DOI: [10.1016/j.jalz.2011.03.008](https://doi.org/10.1016/j.jalz.2011.03.008) (2011).
58. McKeith, I. G. *et al.* Diagnosis and management of dementia with Lewy bodies: third report of the DLB Consortium. *Neurology* **65**, 1863–1872, DOI: [10.1212/01.wnl.0000187889.17253.b1](https://doi.org/10.1212/01.wnl.0000187889.17253.b1) (2005).
59. Mendez, M., Ghajarania, M. & Perryman, K. Posterior cortical atrophy: Clinical characteristics and differences compared to Alzheimer's disease. *Dementia Geriatr. Cogn. Disord.* **14**, 33–40, DOI: [10.1159/000058331](https://doi.org/10.1159/000058331) (2002).
60. Rascovsky, K. *et al.* Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* **134**, 2456–2477, DOI: [10.1093/brain/awr179](https://doi.org/10.1093/brain/awr179) (2011).
61. Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R. & Kievit, R. Raincloud plots: a multi-platform tool for robust data visualization. *PeerJ Prepr.* **6**, e27137v1, DOI: [10.7287/peerj.preprints.27137v1](https://doi.org/10.7287/peerj.preprints.27137v1) (2018).
62. Ashburner, J. A fast diffeomorphic image registration algorithm. *NeuroImage* **38**, 95–113, DOI: [10.1016/j.neuroimage.2007.07.007](https://doi.org/10.1016/j.neuroimage.2007.07.007) (2007).
63. Friston, K. J. *et al.* Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2**, 189–210, DOI: [10.1002/hbm.460020402](https://doi.org/10.1002/hbm.460020402) (1994).
64. Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273–289, DOI: [10.1006/nimg.2001.0978](https://doi.org/10.1006/nimg.2001.0978) (2002).
65. Kivisaari, S. L., Probst, A. & Taylor, K. I. The perirhinal, entorhinal, and parahippocampal cortices and hippocampus: An overview of functional anatomy and protocol for their segmentation in MR Images. In Ulmer, S. & Jansen, O. (eds.) *fMRI*, 239–267, DOI: [10.1007/978-3-642-34342-1_19](https://doi.org/10.1007/978-3-642-34342-1_19) (Springer Berlin Heidelberg, 2013), 2nd edition edn.

Acknowledgements

Research in this publication was supported by Swiss National Science Foundation Ambizione Fellowship #PZ00P1_126493 to KIT, Early Postdoctoral Mobility Fellowship #BBSP_146854 and Academy of Finland grant #286070 to SLK; Professional Staff Congress–City University of New York grant #64535-00-42 and NIH-National Institute of General Medical Sciences of the National Institutes of Health (NIH-NIGMS) under award number SC3-GM122662 to NSF; MRI scan data were supported in part by research grants from GlaxoSmithKline (RES104859 – Basel) and the Novartis Foundation to AUM.

Author contributions statement

NSF, AUM, KIT and SLK conceived the experiment; NSF, SLK and KIT performed analyses and wrote the manuscript; MS and AUM participated in patient data collecting; all authors reviewed the manuscript.

Additional information

Competing interests Authors NSF, SLK and MS declare no competing interests. AUM received fees received from the Advisory Boards of Merz, Vifor, Schwabe, and Roche. The fees were not associated with the submitted work. KIT is a full-time employee of Hoffmann-La Roche Ltd.

Data availability statement

The datasets generated and analyzed in the current study contain clinical patient data. These data cannot be made publicly available due to constraints imposed by the local ethics committee at the time of the data collection.