

# UC Irvine

## UC Irvine Previously Published Works

### Title

Cognitive modeling of the Mnemonic Similarity Task as a digital biomarker for Alzheimer's disease

### Permalink

<https://escholarship.org/uc/item/3ww5n6r5>

### Journal

Alzheimer's & Dementia, 20(10)

### ISSN

1552-5260

### Authors

Vanderlip, Casey R

Lee, Michael D

Stark, Craig EL

### Publication Date

2024-10-01

### DOI

10.1002/alz.14163

Peer reviewed

RESEARCH ARTICLE

# Cognitive modeling of the Mnemonic Similarity Task as a digital biomarker for Alzheimer's disease

Casey R. Vanderlip<sup>1</sup>  | Michael D. Lee<sup>2</sup> | Craig E. L. Stark<sup>1</sup> 

<sup>1</sup>Department of Neurobiology and Behavior, 1424 Biological Sciences III, University of California, Irvine, Irvine, California, USA

<sup>2</sup>Department of Cognitive Science, 3151 Social Sciences Plaza A, University of California, Irvine, Irvine, California, USA

## Correspondence

Craig E. L. Stark, Department of Neurobiology and Behavior, 1424 Biological Sciences III, University of California Irvine, Irvine, CA 92697 USA.

Email: [cestark@uci.edu](mailto:cestark@uci.edu)

## Funding information

Stanford Aging and Memory Study, Grant/Award Numbers: R01 AG066683, P30 AG066519

## Abstract

**BACKGROUND:** The Mnemonic Similarity Task (MST) is a popular memory task designed to assess hippocampal integrity. We assessed whether analyzing MST performance using a multinomial processing tree (MPT) cognitive model could detect individuals with elevated Alzheimer's disease (AD) biomarker status prior to cognitive decline.

**METHOD:** We analyzed MST data from >200 individuals (young, cognitively healthy older adults and individuals with mild cognitive impairment [MCI]), a subset of which also had existing cerebrospinal fluid (CSF) amyloid beta ( $A\beta$ ) and phosphorylated tau (pTau) data using both traditional and model-derived approaches. We assessed how well each could predict age group, memory ability, MCI status,  $A\beta$ , and pTau status using receiver operating characteristic analyses.

**RESULTS:** Both approaches predicted age group membership equally, but MPT-derived metrics exceeded traditional metrics in all other comparisons.

**DISCUSSION:** A MPT model of the MST can detect individuals with AD prior to cognitive decline, making it a potentially useful tool for screening and monitoring older adults during the asymptomatic phase of AD.

## KEYWORDS

Alzheimer's disease, cognitive modeling, digital biomarker, early diagnosis, Mnemonic Similarity Task

## Highlights

- The MST, along with cognitive modeling, identifies individuals with memory deficits and cognitive impairment.
- Cognitive modeling of the MST identifies individuals with increased AD biomarkers prior to changes in cognitive function.
- The MST is a digital biomarker that identifies individuals at high risk of AD.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

## 1 | BACKGROUND

Alzheimer's disease (AD) is marked by a gradual decline in memory and cognitive abilities that are often observed only after amyloid beta (A $\beta$ ) and phosphorylated tau (pTau) are already present.<sup>1-4</sup> Elevated levels of A $\beta$  and pTau increase the risk of cognitive decline,<sup>5,6</sup> making this preclinical stage of AD a critical window for early detection and intervention.<sup>7</sup> During this phase, therapies for AD, including, but not limited to, those targeting A $\beta$  and pTau, could be most effective, prior to irreversible neuronal loss.<sup>8</sup>

Measuring A $\beta$  and pTau is possible using positron emission tomography (PET) and cerebrospinal fluid (CSF), but both invasive and costly, limiting their general application in clinical settings.<sup>9,10</sup> Recent developments in blood testing for A $\beta$  and pTau levels show promise in overcoming these barriers,<sup>11,12</sup> enabling them to become useful clinical tools. The early detection of cognitive impairment via digital biomarkers is also showing promise.<sup>13-16</sup> These non-invasive assessments, which can often be remotely self-administered, could complement blood tests in identifying individuals at future risk of decline, as they may detect different aspects of AD progression. Supporting this, work has found that combining blood biomarkers with cognitive tests offers a more accurate prediction of AD than using either method alone.<sup>17</sup> However, traditional cognitive tests have been less effective in identifying individuals at high risk of AD before cognitive symptoms appear.<sup>18</sup> This underscores the need for refined cognitive tasks that can detect subtle cognitive changes linked to AD pathology and aid in early diagnosis when combined with biomarker analysis.

The Mnemonic Similarity Task (MST) is a promising tool as it is designed to tax hippocampal function through its emphasis on pattern separation, a process central to rapidly learning new, arbitrary information.<sup>19-21</sup> Performance on the pattern separation component of the MST (the Lure Discrimination Index [LDI]) has been associated with functional and structural changes within the hippocampus and related structures while the recognition memory aspect (REC) of the task has not.<sup>22,23</sup> Given that the hippocampus (and entorhinal cortex, which serves as a gateway to the hippocampus) is one of the first affected by aging and AD,<sup>24-27</sup> it is unsurprising that performance declines with age and AD.<sup>28,29</sup> Further, work has demonstrated that the MST can predict early cognitive changes in AD, and this task has been used in multiple clinical trials, including A4 and HOPE4MCI.<sup>30-33</sup>

The MST's traditional metrics are designed to be simple and robust but obscure potentially useful aspects of memory performance. Cognitive modeling of an individual's memory can give a richer understanding of mechanisms<sup>34</sup> and how these are altered by aging or cognitive impairments.<sup>35-37</sup> Recently we developed a multinomial processing tree (MPT) cognitive model to analyze performance on the MST using Bayesian methods that both fit individual participant performance and identified individual differences in memory and response strategies.<sup>38</sup>

Here, we applied this approach to determine whether it aided the MST's ability to discriminate various groups of individuals based on age, cognitive status, and A $\beta$ /pTau status. We found that the MPT model was clearly superior to traditional metrics, particularly

### RESEARCH IN CONTEXT

- 1. Systematic review:** Using traditional sources, such as PubMed, the authors identified studies linking the Mnemonic Similarity Task (MST) to Alzheimer's disease (AD) risks. Most studies have shown that the MST can identify individuals with cognitive impairment, yet the MST's ability to identify older adults with increased AD biomarker status remains underexplored.
- 2. Interpretation:** If digital biomarkers can be developed and validated to reflect AD pathology, they might offer a non-invasive, cost-effective way to predict Alzheimer's risk or monitor disease or treatment progression. We show how a cognitive-modeling-based analysis method not only classifies older adults with changes in memory and cognitive function but also identifies older adults with increased cerebrospinal fluid A $\beta$  and pTau181 prior to clinical changes in cognitive ability.
- 3. Future directions:** We encourage future studies to employ cognitive modeling of the MST as a digital biomarker to identify individuals at high risk for future cognitive decline.

with regard to A $\beta$ /pTau status, highlighting the MST's potential as an effective digital biomarker for early AD detection and monitoring.

## 2 | METHODS

Data from this study came from two previously published works. Analyses 1 to 3 used participants from Stark et al. (2013), while analysis 4 used data from Trelle et al. (2021). Both used the same format as the MST, and both works attempted to identify cognitively "healthy" adults as part of their screening and assessment procedures.

### 2.1 | Analysis 1: Predicting age group from cognitive modeling of MST

For predicting age group, people who were less than 40 years old ( $n = 27$ , age =  $27.41 \pm 5.7$ , 16 females) were classified as young, and those over 60 ( $n = 46$ , age =  $71.33 \pm 6.4$ , 28 females) were considered aged. All individuals were initially screened to be cognitively healthy without impairment using a battery of cognitive tasks. These include the Mini-Mental State Exam (MMSE),<sup>39</sup> Wechsler Memory Scale Logical Memory,<sup>40</sup> Rey Auditory Verbal Learning Test,<sup>41</sup> Verbal Fluency,<sup>42</sup> Digit Span,<sup>43</sup> Trails A and B,<sup>44</sup> and Letter Number Sequencing,<sup>43</sup> and the Wechsler Adult Intelligence Scale III.<sup>43</sup> All individuals scored within 1.5 standard deviations of the mean of their age group for all neuropsychological measures.

## 2.2 | Analysis 2: Predicting memory deficits in older adults using cognitive modeling of the MST

Significant work has used the Rey Auditory Verbal Learning Test (RAVLT) to differentiate older adults into separate groups based on cognitive function.<sup>45,46</sup> The RAVLT consists of learning a list of 15 words and recalling them after a delay of 15 min, and the delay score ranges from 0 to 15 and reflects the number of words correctly recalled after the delay. In the original report, older adults were split into thirds based on their RAVLT performance to parallel work in rodents that examined aged unimpaired (AU) and aged impaired (AI) groups.<sup>29</sup> It is important to note that AI individuals (RAVLT of 5 to 8) are still within their age-based norms and are not clinically impaired. AU individuals (RAVLT of 12 to 15) have a performance similar to young adults (this threshold is often used as part of the "SuperAger" criteria). However, here we used a threshold of 9 on the RAVLT to split older adults into either individuals with age-related memory deficits (AMD) or no age-related memory deficits (NMD) since performance above 9 is above age-matched norms.<sup>47</sup> Similar to prior work, individuals who scored higher than 9 were considered NMD ( $n = 31$ , age =  $71.29 \pm 6.79$ , 18 females), and those who scored 9 or below, but within normal limits of their age group, were considered AMD ( $n = 15$ , age =  $71.40 \pm 5.8$ , 10 females).<sup>46-49</sup>

## 2.3 | Analysis 3: Predicting cognitive status in cognitively older adults

To predict whether older adults were cognitively normal (CN) or had mild cognitive impairment (MCI) using the MST, the same 46 adults over the age of 60 from the previous two analyses were used for older adults who are cognitively intact ( $n = 46$ , age =  $71.33 \pm 6.4$ , 28 females). A further 10 individuals (age =  $76.30 \pm 6.78$ , 5 females) who were diagnosed with amnesic MCI were also included. Individuals with MCI were diagnosed by the University of California, Irvine (UCI) Alzheimer's Disease Research Center (ADRC). All individuals with MCI had a global Clinical Dementia Rating (CDR) of 0.5, a memory complaint, and impaired memory function on neuropsychological testing. A final diagnosis of amnesic MCI was reached by neurologists and neuropsychologists at clinical consensus conferences within the UCI ADRC. No participants had a history of neurological or psychiatric disorders, head trauma with loss of consciousness, drug abuse, or dependency.

## 2.4 | Analysis 4: Predicting biomarkers of AD in CN older adults

To predict AD biomarker status, we used previously published data,<sup>50</sup> collected as part of the Stanford Aging and Memory Study (SAMS). One hundred thirty-three older adults (age =  $68.8 \pm 5.8$ , 83 females) were administered the MST and underwent a lumbar puncture to quantify AD biomarkers. All individuals had normal or corrected-to-normal

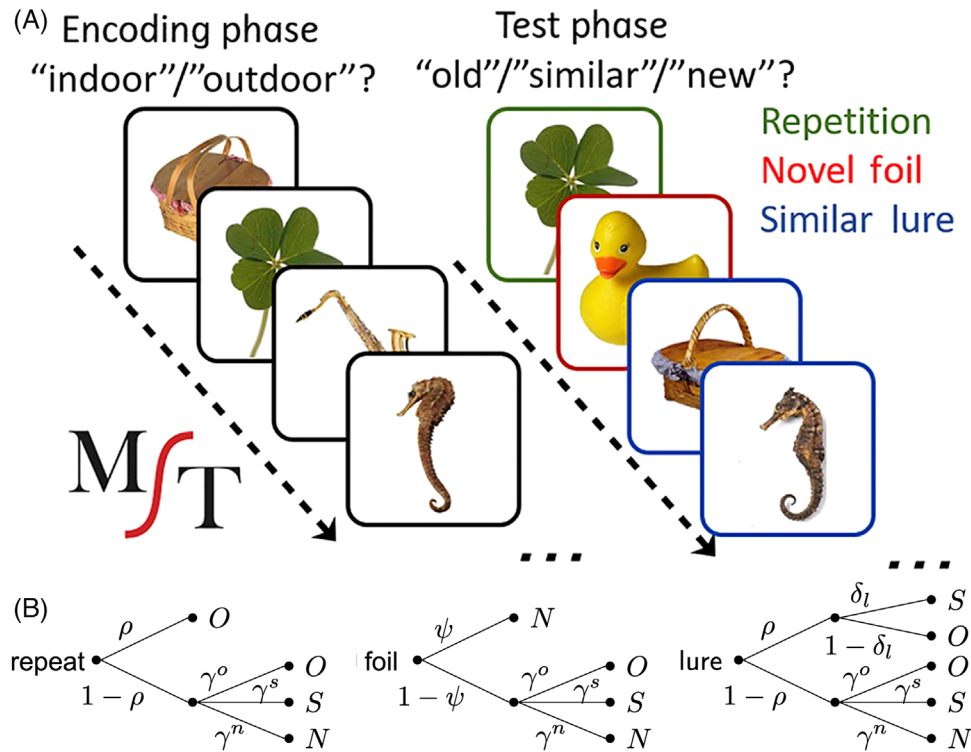
vision/hearing, were right-handed, were native English speakers, and had no history of neurologic or psychiatric disease. Further, each participant had a global CDR score of zero and performance within the normal range on a standardized neuropsychological test battery. Lastly, all participants were deemed CN during a clinical consensus meeting consisting of neurologists and neuropsychologists. The previously derived  $A\beta$  and pTau181 statuses were used in the present analyses (see Trelle et al.<sup>50</sup> for details). Briefly, individuals with an  $A\beta_{42}/A\beta_{40}$  ratio less than 0.0752 ( $n = 34$ , age =  $70.33 \pm 6.04$ , 21 females) were considered A+, and those with an  $A\beta_{42}/A\beta_{40}$  ratio less than 0.0752 ( $n = 99$ , age =  $68.13 \pm 5.60$ , 63 females) were A-. Individuals with pTau181 greater than 42 pg/mL were considered T+ ( $n = 31$ , age =  $71.86 \pm 6.14$ , 21 females) with the remaining considered T- ( $n = 102$ , age =  $67.73 \pm 5.32$ , 63 females).

The MST is a widely used cognitive task, that is, thought to critically tax hippocampal pattern separation (Figure 1A).<sup>23,29</sup> Both data sources used the traditional version of the MST, which consists of an incidental encoding phase and an explicit test phase. During the encoding phase, individuals made successive indoor/outdoor judgments for 128 images (2 s each, 0.5 ISI, color objects on a white background) via a button press. Immediately following the encoding phase, participants were given instructions for a recognition memory test, where they were told to identify objects as either "old" (the same picture as before), "similar" (indicating this is similar but not identical to the studied item, eg, a different exemplar, a rotation), or "new" via a button press. Here, participants saw 192 images (2 s each, 0.5 inter-stimulus-interval (ISI)) and responded to each of these images. Images consisted of 64 exact repeats from the encoding phase (targets), 64 completely novel images (foils), and 64 images that were similar but not identical to images seen during encoding (lures).

Multiple behavioral metrics were extracted from the MST (Table 1), including the traditional REC and LDI scores. REC is a commonly used measure of recognition memory test and is the probability of "old" responses given to the target items minus the corresponding probability of "old" responses given to the foils (to correct for response bias). To quantify the ability to discriminate between similar lures, the LDI is the difference between the probability of giving a "similar" response to lure items and the probability of giving a "similar" response to the foils to account for any bias individuals may have in using the "similar" response overall. For a follow-up analysis, we also quantified the rate of "old" responses for target trials (hits), the rate of "similar" responses for lure trials (correct rejections of lures), and the rate of "new" responses for foil trials (correct rejections of foils). Further, we attempted to obtain a readout of guessing by calculating the rate of "old" responses on foil trials, the rate of "similar" responses on foil trials, and the rate of "new" responses on target trials.

## 2.5 | Cognitive modeling

Cognitive modeling provides a useful tool for inferring latent psychological variables beyond traditional measurements. Previously, we used cognitive modeling to model subject-level performance on the MST



**FIGURE 1** Cognitive modeling of MST. (A) Sample stimuli used during incidental encoding phase and subsequent old/similar/new recognition task. (B) Probability tree diagrams of MPT model, demonstrating decision-making process utilized within old/similar/new version of MST. Responses old (O), similar (S), and new (N) are modeled to be due to a decision tree. For example, on a repeat trial, the model assumes there is some probability,  $\rho$ , that at least a gist-level successful retrieval is made. If that occurs, the modeled participant will respond "old." If that fails, the modeled participant will respond "old," "similar," or "new" based on the guessing/bias parameters  $\gamma$ . MST, Mnemonic Similarity Task.

in young adults<sup>38</sup> using the MPT framework, a common approach for cognitive modeling of REC tasks. The MPT framework assumes that cognitive processes can be divided into discrete categories or decision points (Figure 1B). Briefly, when a repeated item appears, we assume there is a probability ( $\rho$ ) that the item is successfully matched with memory in at least a basic gist or "familiarity" form, leading to an "old" response. Failing that, we assume that a guess is made with unique probabilities (response biases) for each of the three responses. Similarly, when an unrelated foil is present, there is a probability ( $\psi$ ) that the lack of a match to memory is sufficiently clear that a "no" response is made; failing that, a three-choice guess is made.

When a similar lure is presented, there is an initial decision point involving recognizing some degree of match between the object presented and the memory of one previously encountered, based on the same  $\rho$  as above. This level of match is modeled to reflect a simpler item-, gist-, or familiarity-based match (for both lures and repeated items). If this is unsuccessful, a three-choice guess happens, as before. If successful, there is a second decision point based on a set of similarity-based probabilities ( $\delta$ ) capturing whether the memory retrieval contains the richer details required to reject the item as only being similar to the studied item. If successful, a "similar" response is made; if it is unsuccessful, an "old" response is made.

Posterior distributions for metrics within MPT models were estimated at the subject level from trial-by-trial (condition (repeat, lure, foil), response (old, similar, new), and lure bin (one to five)) experimen-

tal data using Just Another Gibbs Sampler (JAGS), which is a program designed for Bayesian hierarchical models.<sup>51</sup> Models were run with eight separate Markov chain Monte Carlo chains with 1000 samples per chain after removing 1000 burn-in samples. We used posterior means as point estimates for multiple metrics of interest (Table 1). These metrics include  $\rho$ , which reflects the probability of remembering items,  $\lambda$ , based on  $\delta$  and designed to capture the ability to discriminate remembered items from lures,  $\psi$ , the probability of remembering that an item was not studied,  $\gamma^o$  (probability of guessing "old"),  $\gamma^N$  (probability of guessing "new"), and  $\gamma^s$  (probability of guessing "similar"). Supplemental information and code to run the model are provided at <https://github.com/mdlee/mpt4mst>.

## 2.6 | Statistical analyses

All analyses were done in Python. Logistic regressions were run using statsmodels<sup>52</sup> to predict age group, clinical status, and biomarker status, for example, from various sets of metrics. Areas under the curve (AUC) measures were derived from receiver operating characteristic (ROC) curves of the logistic regressions. To compare model fits, we calculated the Bayesian information criterion (BIC) of each model.<sup>53</sup> Absolute differences in BICs of greater than 2 were considered reliable. Considering metrics in isolation and in combination with other metrics from the same approach does allow for direct comparisons

**TABLE 1** Descriptions of both traditional and model-based metrics on the MST.

Metric	Type	Definition
REC	Traditional	Recognition memory score
LDI	Traditional	Reflects ability to discriminate between similar lures
p(Old Repeat)	Traditional	Probability of responding "old" for repeats
p(Sim Lure)	Traditional	Probability of responding "similar" to lures
p(New Foil)	Traditional	Probability of responding "new" for foils
p(Old Foil)	Traditional	Probability of responding "old" for foils
p(Sim Foil)	Traditional	Probability of responding "similar" for foils
p(New Repeat)	Traditional	Probability of responding "new" for repeats
$\rho$	Modeled	Probability of remembering items at a gist level
$\lambda$	Modeled	Ability to discriminate remembered items from lures
$\psi$	Modeled	Probability of remembering that an item was not studied
$\gamma^O$	Modeled	Probability of guessing "old"/bias to respond "old"
$\gamma^S$	Modeled	Probability of guessing "similar"/bias to respond "similar"
$\gamma^N$	Modeled	Probability of guessing "new"/bias to respond "new"

Abbreviations: LDI, lure discrimination index; REC, recognition memory.

across the techniques. However, the metrics are not independent of each other, and the two approaches differ in the number of variables considered. Therefore, to identify how each variable acts in conjunction with the others, we performed an 8-choose-4 combinatorial analysis and quantified the number of times each metric appeared in the top third of AUCs from 8-choose-4 analyses. Independent sample *t* tests were used to examine group differences in traditional and model-based metrics.<sup>54</sup> To investigate group changes in guessing strategies, Kolmogorov–Smirnov tests were used because data were proportioned and therefore not normally distributed. For all analyses,  $p < 0.05$  was considered reliable.

### 3 | RESULTS

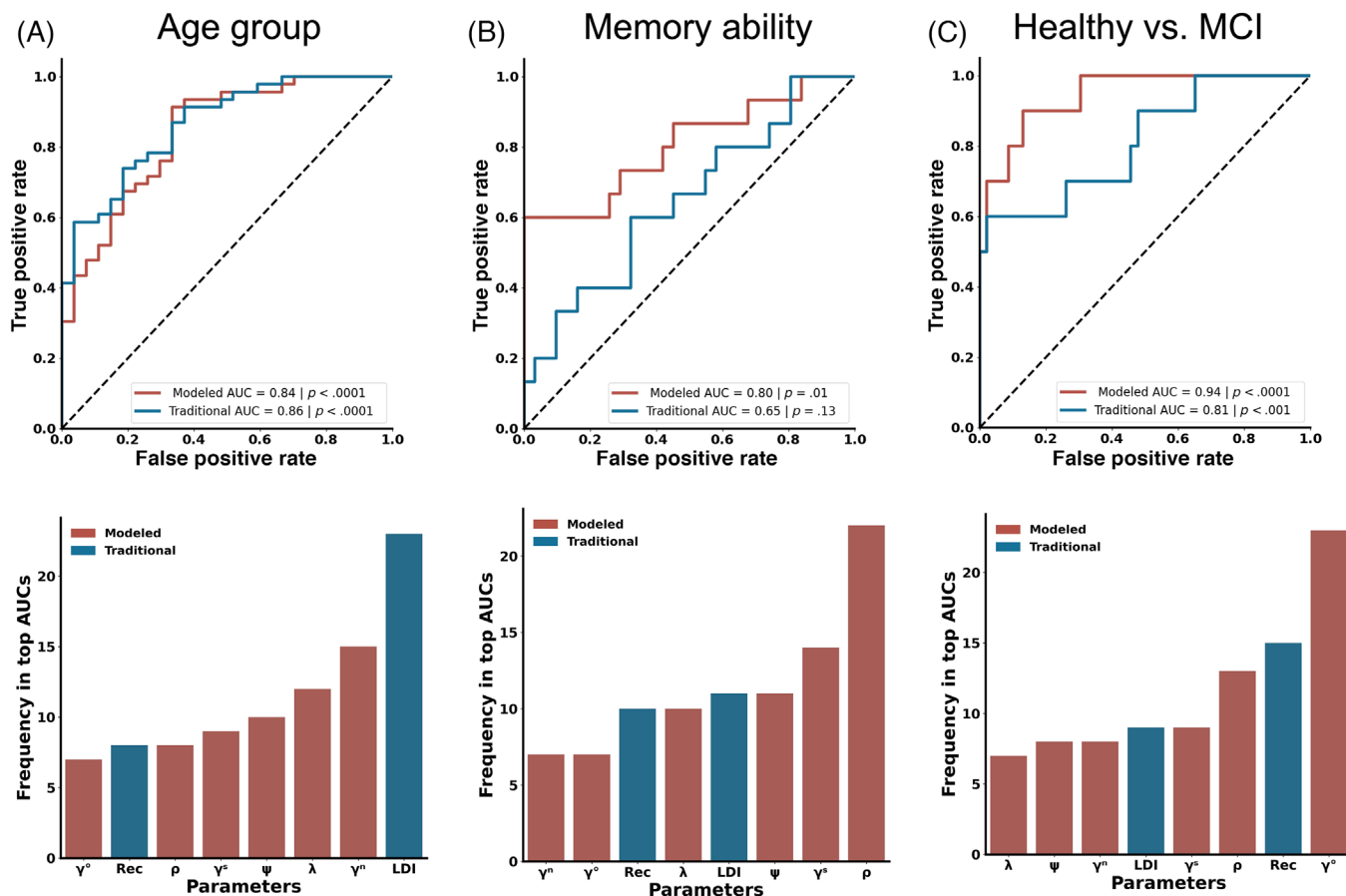
To investigate if the Bayesian MPT cognitive model could adequately fit the observed data, we conducted posterior predictive checks, which are the standard Bayesian approach to testing the ability of the model to redescribe the data on which parameter inferences are based, using the approach developed in Lee and Stark (2023). These checks demonstrated that the model was able to accurately redescribe the empirical data. We plotted these posterior predictive checks for each participant from Stark et al., 2013 (Figure S1A) and Trelle et al., 2021 (Figure S1B). Also, the measures from the MPT model were precise with the width of the confidence intervals for all the measures below 0.5. Previously, we demonstrated that the traditional REC correlated with  $\rho$ , while LDI correlated with  $\lambda$  (previously denoted by  $\tau$ ).<sup>37</sup> This is because REC and  $\rho$  are both measures thought to reflect REC on a task, while LDI and  $\lambda$  are thought to be related to mnemonic discrimination. Our first goal was to assess the relationship between the traditional and modeled metrics in the two datasets.<sup>29,50</sup> As in the prior work, we found strong correlations between these variables in both datasets (Stark et al., 2013; REC

vs  $\rho$ :  $r = 0.73$ , LDI vs  $\lambda$ :  $r = -0.90$ , Trelle et al., 2021; REC vs  $\rho$ :  $r = 0.77$ , LDI vs  $\lambda$ :  $r = -0.90$ ). A full correlation matrix for both Stark et al. (Figure S2A) and Trelle et al. (Figure S2B) revealed similar correlations between datasets. These results demonstrate that the model-based metrics derived here are similar to prior findings. With this, we performed four key analyses, using two datasets, to assess if traditional or model-based metrics were superior at identifying individuals at risk for AD.

#### 3.1 | Analysis 1: Traditional metrics and model-based metrics of MST equally predict age group status

Given that extensive work has demonstrated that older adults are impaired on the MST, we assessed whether cognitive modeling could enhance the ability to differentiate between younger and older adults.<sup>23,29</sup> Considered individually, there was no reliable difference in REC between age groups, while LDI was significantly lower in older adults (Figure S3A, REC:  $t(71) = 1.19$ ,  $p = 0.28$ , Figure S3B, LDI:  $t(71) = 5.71$ ,  $p < 0.0001$ ). When examining modeled metrics individually,  $\rho$  showed no reliable age differences, while  $\psi$  and  $\lambda$  were lower in older compared to younger adults (Figure S3C,  $\rho$ :  $t(71) = 0.22$ ,  $p = 0.83$ , Figure S3D,  $\psi$ :  $t(71) = 2.62$ ,  $p < 0.05$ , Figure S3E,  $\lambda$ :  $t(71) = -5.69$ ,  $p < 0.0001$ ). A multiple logistic regression using the traditional LDI and REC as predictors achieved an AUC of 0.86 (Figure 2A,  $p < 0.0001$ ). Model-based metrics, with  $\rho$ ,  $\psi$ ,  $\tau$ , and guessing strategies ( $\gamma^O$ ,  $\gamma^N$ , and  $\gamma^S$ ) as predictors, yielded a similar AUC of 0.84 (Figure 2A, top,  $p < 0.0001$ ), suggesting that model-based metrics did not outperform traditional metrics in predicting age group.

To appreciate better the impact each variable might have in conjunction with the others, we performed an 8-choose-4 combinatorial



**FIGURE 2** Comparison of traditional (blue) to MPT model (red) performance in questions 1 to 3. ROC curves (top) demonstrating predictive value of traditional and model-based measures and frequency of presence in top 30 AUCs in 8-choose-4 combinatorial analyses (bottom) are shown. (A) Comparison of age group predictions showing no significant difference between traditional measures and MPT model-based measures. LDI emerges as the most frequent metric in the top third of AUCs. (B) Metrics derived from MPT model better predict the performance of high-versus low-performing older adults. The gist-based recognition memory signal metric ( $\rho$ ) is the predominant metric in the 8-choose-4 analysis for predictive accuracy. (C) MPT model-based metrics were superior at identifying healthy versus MCI. Within the 8-choose-4 analysis,  $\gamma^o$  is the leading metric for MCI prediction. AUC, area under a curve; LDI, lure discrimination index; MCI, mild cognitive impairment; MPT, multinomial processing tree; MST, Mnemonic Similarity Task.

analysis and identified how often each factor occurred in the top third of the resulting AUCs. This revealed that the LDI was the most common metric in distinguishing younger and older adults, appearing in virtually all the top-performing models and almost twice as often as the most frequent MPT model-based metric (Figure 2A, bottom). Thus, when considering the simpler task of predicting age group membership, we found no evidence that the MPT model was superior to the traditional approach.

### 3.2 | Analysis 2: Model-based metrics better identify memory ability of older adults

Aging is often associated with substantial heterogeneity, with a subset of older adults performing at higher levels on neuropsychological memory assessments compared to their peers. Therefore, we next asked if performance on the MST along with cognitive modeling could aid in dissociating across levels of cognitive function in healthy adults by

discriminating NMD versus AMD. Considering each variable individually, REC and LDI levels were similar in NMD and AMD (Figure S3F, REC:  $t(44) = 0.22$ ,  $p = 0.22$ , Figure S3G, LDI:  $t(44) = 1.72$ ,  $p = 0.09$ ). When measuring model-based metrics,  $\rho$  was significantly higher in NMD compared to AMD individuals with no difference in  $\psi$  or  $\lambda$  (Figure S3H,  $\rho$ :  $t(44) = 3.10$ ,  $p < 0.01$ ; Figure S3I,  $\psi$ :  $t(44) = 1.248$ ,  $p = 0.22$ ; Figure S3J,  $\lambda$ :  $t(44) = -0.31$ ,  $p = 0.76$ ). When LDI and REC were combined, a multiple logistic regression did not successfully distinguish NMD versus AMD (AUC = 0.65,  $p = 0.13$ , Figure 2B). However, multiple logistic regression with model-based metrics was able to stratify NMD from AMD with an AUC of 0.80 ( $p < 0.05$ ). When assessing combinations of traditional and model-based metrics in an 8-choose-4 combinatorial analysis,  $\rho$  emerged as the most consistent metric in the top-performing models, with LDI appearing as a distant fourth most consistent (Figure 2B, bottom). This suggests that the MPT model provides a more accurate identification of memory ability in older adults than traditional metrics but that this is driven heavily by the model's estimate of how well individuals remember at least the gist of an item.

### 3.3 | Analysis 3: Model-based metrics better predict MCI status

We next investigated whether cognitive modeling of the MST could better identify individuals with MCI compared to traditional metrics. We found that individuals with MCI had significantly lower REC performance compared to CN older adults, but there were no differences between groups in LDI scores (Figure S3K, REC:  $t(52) = 4.73$ ,  $p < 0.0001$ ; Figure S3L, LDI:  $t(52) = 0.77$ ,  $p = 0.44$ ). We also found that  $\rho$  decreased in individuals with MCI, but no difference in groups for  $\psi$  or  $\lambda$  (Figure S3M,  $\rho$ :  $t(52) = 5.51$ ,  $p < 0.0001$ ; Figure S3N,  $\psi$ :  $t(52) = 0.99$ ,  $p = 0.33$ ; Figure S3O,  $\lambda$ :  $t(52) = -0.16$ ,  $p = 0.87$ ). In the multiple logistic regression, we found that the combination of REC and LDI could classify MCI status with good accuracy (AUC = 0.81,  $p < 0.001$ , Figure 2C). However, cognitive-model-based metrics offered superior predictive power, achieving an AUC of 0.94 ( $p < 0.0001$ ). Permutation analysis found that  $\gamma^O$  was the most influential metric, appearing in all the top third of models (Figure 2C, bottom). This suggests that the MPT model is superior at detecting MCI over traditional metrics largely due to the ability to derive differences in guessing strategy on the task.

### 3.4 | Analysis 4: Model-based metrics can better predict A $\beta$ and Tau status in CN older adults

We next evaluated whether the MPT model of the MST could detect A $\beta$  status in cognitively healthy older adults, classified as A $\beta$ + or A $\beta$ - via CSF A $\beta$ 42/A $\beta$ 40 ratios. A $\beta$ + individuals had decreased REC scores but equivalent LDI performance compared to A $\beta$ - counterparts (Figure S4A, REC:  $t(131) = 2.68$ ,  $p < 0.01$ , Figure S4B, LDI:  $t(131) = 0.33$ ,  $p = 0.74$ ). Further,  $\rho$  was lower in A $\beta$ + compared to A $\beta$ - older adults with no group differences in  $\psi$  and  $\lambda$  (Figure S4C,  $\rho$ :  $t(131) = 2.54$ ,  $p < 0.05$ , Figure S4D,  $\psi$ :  $t(131) = 1.11$ ,  $p = 0.27$ , Figure S4E,  $\lambda$ :  $t(131) = -0.53$ ,  $p = 0.60$ ). Multiple logistic regression with traditional metrics could modestly predict amyloid status (AUC = 0.64,  $p < 0.05$ , Figure 3A). On the other hand, multiple logistic regression with model-based metrics better predicted amyloid status (AUC = 0.73,  $p < 0.05$ ). When conducting an 8-choose-4 combinatorial analysis to investigate the impact each variable might have in relation to the others,  $\gamma^O$  was the most predictive metric among the top third of AUCs. Interestingly,  $\gamma^O$  was represented in nearly all the top models and twice as often as both traditional metrics (Figure 3A, bottom). The model thus better identifies asymptomatic individuals with elevated amyloid burden due to its ability to derive differences in guessing "old."

While both A $\beta$  and pTau are biomarkers for AD, pTau has a stronger link to cognitive decline and may better predict disease progression. Somewhat surprisingly, CN older adults with elevated pTau levels did not differ on either traditional or model-based metrics (REC, LDI,  $\rho$ ,  $\psi$  and  $\lambda$ ) compared to those with normal pTau levels (Figure S4F-J, all  $ps > .10$ ). Likewise, a multiple logistic regression with REC and LDI failed to predict pTau status (AUC = 0.50,  $p = 0.91$ , Figure 3B). Importantly, the logistic regression with the model-based metrics did predict pTau status (AUC = 0.71,  $p < 0.05$ ). When conducting an

8-choose-4 combinatorial analysis, the single clearly most reliable metric was  $\gamma^O$ , appearing more than twice as much as the next most important metric ( $\psi$ ) (Figure 3B, bottom). Further, every metric from the MPT model was more represented than REC and LDI in the top third of models. Overall, the cognitive model outperformed traditional metrics in predictive accuracy, suggesting its effectiveness in early AD screening.

### 3.5 | Changes in model-derived guessing strategies with age, cognitive impairment, and biomarker status

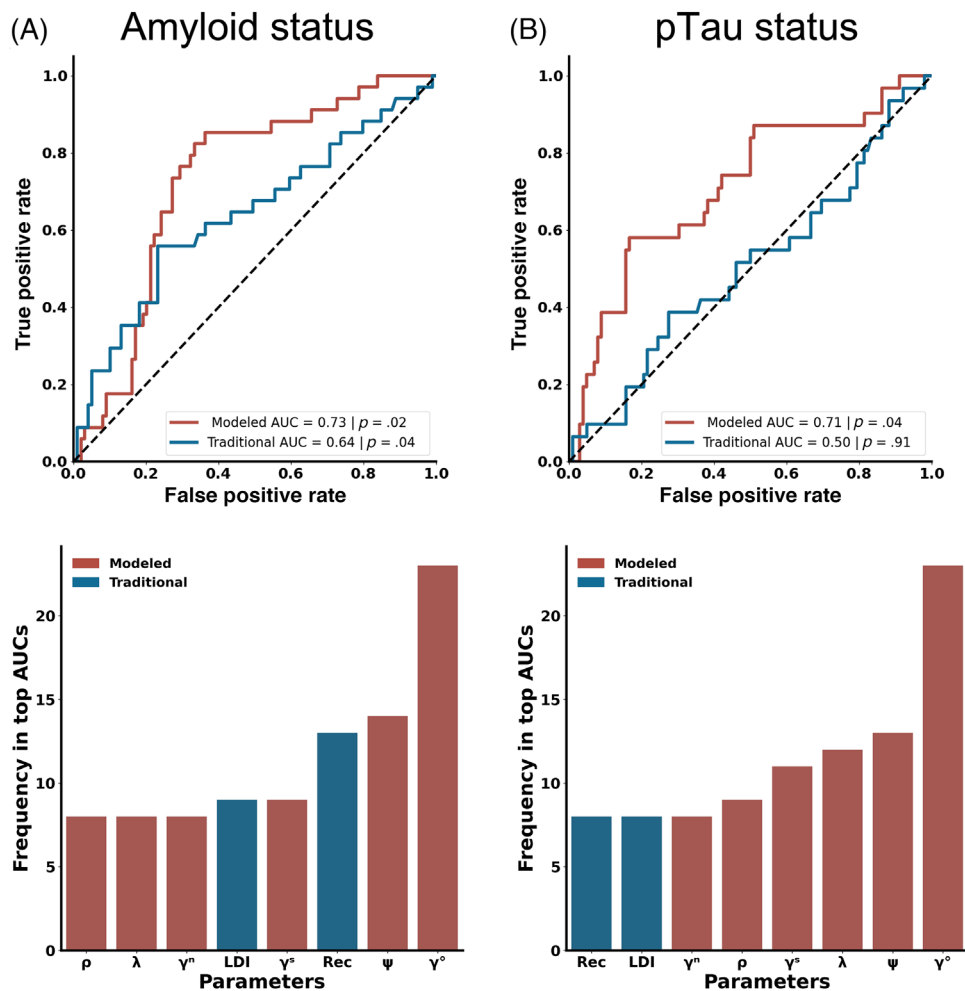
Given that modeling guessing probabilities was informative for predicting AD biomarker status and cognitive status, we took a deeper dive into guessing strategies. We first explored how model-based guessing strategies changed across age groups. Cognitive modeling suggested that younger adults tended to guess "similar" more frequently and "new" less frequently than older adults (Figure 4A, Kolmogorov-Smirnov test;  $\gamma^S$ :  $D = 0.37$ ,  $p < 0.05$ ;  $\gamma^N$ :  $D = 0.38$ ,  $p < 0.05$ ), without any significant age-related differences for guessing "old" ( $\gamma^O$ :  $D = 0.25$ ,  $p = 0.20$ ). This pattern suggests an age-related shift from guessing "similar" to "new." Interestingly, however, no significant differences in guessing strategies were found between healthy older adults with and without memory deficits (Figure 4B, Kolmogorov-Smirnov test: all  $p > 0.10$ ). In contrast, older adults with MCI were more inclined to guess "old" and less likely to guess "similar" or "new" compared to cognitively healthy older adults (Figure 4C, Kolmogorov-Smirnov test;  $\gamma^O$ :  $D = 0.6318$ ,  $p < 0.01$ ;  $\gamma^N$ :  $D = 0.4910$ ,  $p < 0.05$ ;  $\gamma^S$ :  $D = 0.5318$ ,  $p < 0.05$ ). These results underscore that aging and MCI distinctly affect guessing strategies.

We next investigated how guessing strategies on the MST varied with AD biomarker status in older adults. Cognitive modeling suggested that those with elevated amyloid displayed a tendency to guess "old" more frequently than their counterparts without elevated amyloid, but this failed to reach significance. Further, there were no significant differences in biases toward guessing "similar" or "new" (Figure 4D, Kolmogorov-Smirnov test;  $\gamma^O$ :  $D = 0.24$ ,  $p = 0.08$ ,  $\gamma^S$ :  $D = 0.21$ ,  $p = 0.18$ ,  $\gamma^N$ :  $D = 0.13$ ,  $p = 0.75$ ). We next investigated whether guessing strategies changed as a function of pTau status. We observed that individuals with elevated pTau levels were more likely to guess "old" and less likely to guess "similar," with no change in the likelihood of guessing "new" (Figure 4E, Kolmogorov-Smirnov test;  $\gamma^O$ :  $D = 0.32$ ,  $p < 0.05$ ;  $\gamma^S$ :  $D = 0.34$ ,  $p < 0.05$ ,  $\gamma^N$ :  $D = 0.19$ ,  $p = 0.34$ ). These differences in predicted guessing strategies further highlight the benefits of the MPT cognitive model of the MST.

### 3.6 | Addition of raw metrics of guessing does not match model-based metrics

Given the differences in guessing across performance ability, impairment level, amyloid, and pTau status in older adults, we next asked whether we could derive guessing strategies based on response

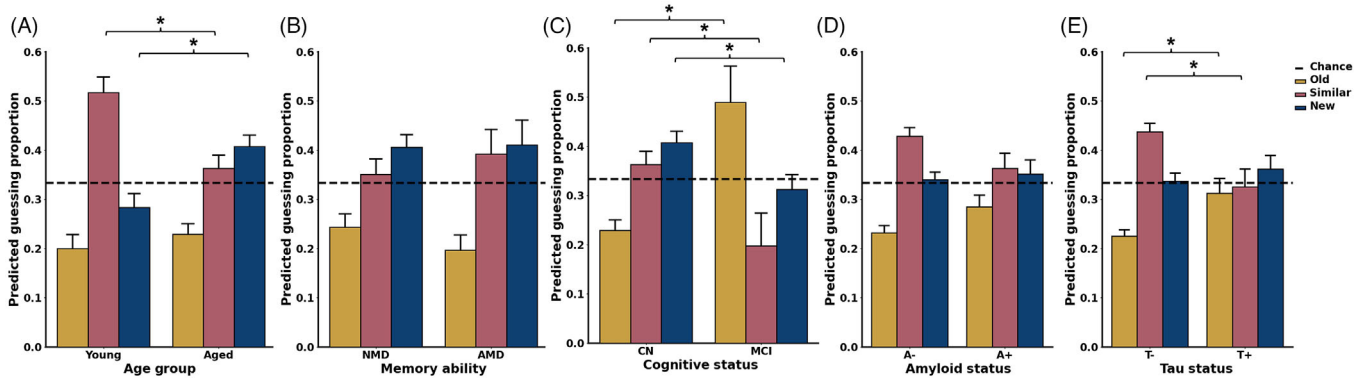




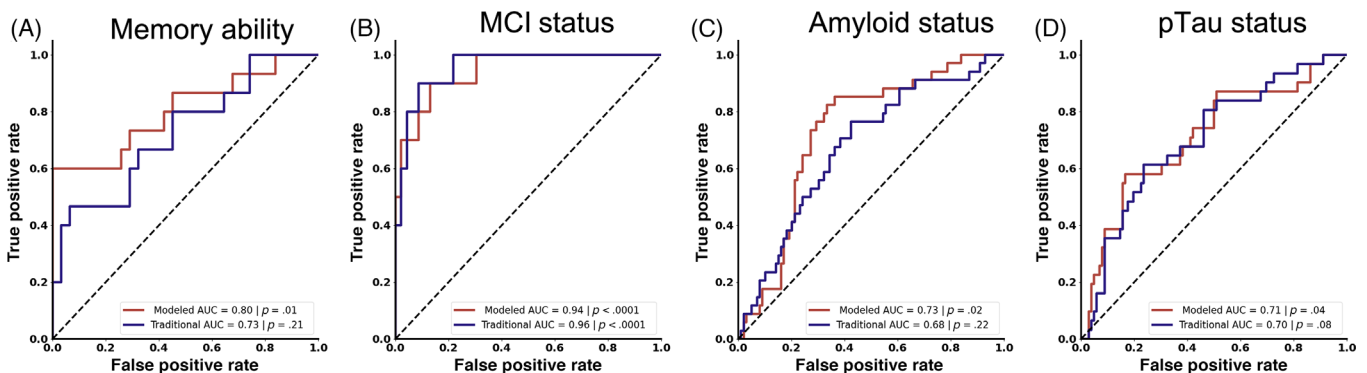
**FIGURE 3** Utilization of the MPT model of the MST for predicting AD biomarker status. (A) Both traditional and MPT model-derived metrics can predict  $A\beta$  status, with the MPT model-based metrics showing superior predictive accuracy. The metric  $\gamma^o$  is identified as the most frequently occurring metric in the top results from  $n$ -choose-four analyses for predicting  $A\beta$  status. (B) The MPT model-based measures, but not traditional measures, can successfully predict pTau status. The metric  $\gamma^o$  is again highlighted as the most common metric in the top 30 AUCs from an  $n$ -choose-four analysis for predicting pTau status. For panels A and B, Traditional measures are in blue, while the MPT model-based metrics are in red.  $A\beta$ , amyloid beta; MPT, multinomial processing tree; MST; Mnemonic Similarity Task; pTau, phosphorylated tau.

patterns. Specifically, we used the proportion of trials an individual responded “old” on foils as a measure of guessing old, the proportion of trials an individual responded “similar” on foils as a measure of guessing similar, and the proportion of trials an individual responded “new” on repeats as a measure of guessing new. REC and LDI incorporate these metrics in their calculations as part of their difference scores, intended to factor out differences in guessing rates. Any baseline shift in the probability of guessing “old” or “similar” would presumably affect both components of the difference metrics, removing what the model-based analyses suggest could be highly informative. Therefore, we also added the raw proportion of trials people responded “new” for foil trials, the proportion of trials an individual responded “old” on repeats, and the proportion of trials an individual responded “similar” on lure trials. Using these six new metrics, we asked whether these metrics could increase the predictive value of the MST to the same level as that provided by the MPT model.

We first asked whether these new behavioral metrics could predict whether CN older adults exhibited memory deficits. In multiple logistic regression, we found that while the raw AUC appeared elevated, this model could still not reliably predict performance level ( $AUC = 0.78$ ,  $p = 0.21$ , Figure 5A). We also investigated whether these differing models better fit the data by examining their respective BIC, which is a metric that reflects the goodness of fit of a model.<sup>55</sup> Using this measure, we found that the MPT model-based metrics better fit the data compared to the probability-based traditional metrics (MPT model BIC: 76.49, raw traditional BIC: 81.86), suggesting that, despite the addition of the raw traditional metrics, cognitive modeling better predicts memory ability. Next, when assessing whether these new probability-based metrics helped with the prediction of MCI status, we found that a multiple logistic regression with these raw traditional metrics significantly bolstered the predictive accuracy of the MST, yielding an AUC of 0.95 ( $p < 0.0001$ , Figure 5B), which aligns with the predictive



**FIGURE 4** Influence of age, memory ability, cognitive status, and AD biomarker status on modeled guessing strategies in MST. (A) Compared to younger adults, older adults exhibit a higher likelihood of guessing “new” and a lower tendency to guess “similar.” (B) Among older adults, those with no age-related memory deficits show no significant differences in guessing strategies when compared to individuals with age-related memory deficits. (C) Individuals with MCI demonstrate a greater bias toward guessing “old” and are less inclined to guess “similar” or “new” relative to cognitively healthy older adults. (D) There are no differences in guessing strategies between  $A\beta^-$  and  $A\beta^+$  older adults. (E) Older adults with elevated pTau levels display a marked shift from guessing similar toward guessing old compared to pTau-negative older adults. For panels A–E, the metric for guessing “old” ( $\gamma^O$ ) is represented in yellow, for guessing “similar” ( $\gamma^S$ ) in pink, and for guessing “new” ( $\gamma^N$ ) in blue.  $A\beta$ , amyloid beta; AD, Alzheimer’s disease; MCI, mild cognitive impairment; MST, Mnemonic Similarity Task; pTau, phosphorylated tau.



**FIGURE 5** Additional raw-derived measures may improve upon traditional metrics but do not match the MPT model-based metrics. (A) Adding raw metrics of guessing fails to predict memory ability, in contrast to cognitive modeling measures. (B) In cases of overt cognitive impairment, such as MCI, raw-derived guessing measures show predictive value comparable to that of the MPT model. (C) Additional raw metrics fail to match the MPT model in predicting  $A\beta$  status. (D) While the additional raw measures improve the predictive ability of traditional approaches for pTau status, their performance does not exceed chance levels. For panels A–D, traditional measures are depicted in purple, and the MPT model-based metrics are illustrated in red.  $A\beta$ , amyloid beta; MCI, mild cognitive impairment; MPT, multinomial processing tree; pTau, phosphorylated tau.

strength of cognitive modeling and actually provided a better fit of the data than model-based metrics (MPT model BIC: 49.77, raw traditional BIC: 45.12).

We next evaluated whether raw traditional metrics could match cognitive modeling in predicting amyloid status. A multiple logistic regression with these metrics did not match the predictive capacity of model-based metrics ( $AUC = 0.68$ ,  $p = 0.22$ , Figure 5C). Moreover, when comparing model fits using the BICs, model-based metrics better fit the data (MPT model BIC: 167.18, raw traditional BIC: 177.12). We next asked if the new probability-based metrics could predict pTau status. These new metrics did show a qualitative improvement in predicting pTau status, but this was not statistically reliable ( $AUC = 0.70$ ,  $p = 0.08$ , Figure 5D) Further, model-based metrics better fit the data compared to probability-based metrics (MPT model BIC: 161.95, raw

traditional BIC: 167.52), reinforcing the superiority of the MPT model in predicting amyloid and pTau status.

## 4 | DISCUSSION

The MST is a widely used memory test that assesses changes in hippocampal integrity in various conditions including age-related cognitive decline and AD.<sup>23</sup> Given that this task is resistant to practice effects and can be easily performed remotely, it has emerged as an ideal candidate for clinical use as a digital biomarker for AD pathology, but direct tests of its validity are needed. We previously demonstrated that a MPT model could be applied to the MST, but we did not know whether this model would aid the predictive value of the MST. In this

study, we compared the predictive value of traditional metrics versus model-based metrics for predicting age group, MCI, and AD biomarkers. We found that the MPT model enhanced the ability of the MST to identify older adults at risk of developing AD prior to cognitive decline, demonstrating that the MST was well suited to enhance early diagnosis of AD.

#### 4.1 | Cognitive modeling of MST identifies differing cognitive capacities in older adults

A large body of work has demonstrated that advancing age is associated with significant impairment on the LDI metric of the MST, while REC remains stable with age.<sup>29,56,57</sup> Similarly, we found that traditional behavioral metrics of the MST predicted age group with high proficiency (AUC = 0.85), and the MPT model did not increase the high predictive value of the MST. In permutation analyses, the LDI was most prominent, further supporting the hypothesis that age-related impairments are due in part to deficits in hippocampal pattern separation.

While age-related impairments are seen on many cognitive tests, there is typically significant heterogeneity within the aging population, with a subset of healthy older adults exhibiting AMD and others showing young-like performance (eg, "SuperAgers" or "aged unimpaired"). Our analyses showed that traditional measures of MST performance did not readily distinguish these two but that the cognitive modeling approach could. This suggests that the MPT model may be more sensitive to differing memory abilities in older adults.

MCI, often considered a precursor to or risk factor for AD, is frequently missed, with perhaps only ~8% of those affected accurately identified.<sup>58</sup> We demonstrated that, like previous work, performance on the MST predicted MCI with an AUC of 0.81.<sup>31,32</sup> However, cognitive modeling significantly improved the predictive accuracy, reaching an AUC of 0.94. We note that multiple mechanisms can cause MCI. While the diagnoses of MCI followed the criterion for MCI due to AD,<sup>59</sup> other mechanisms could have caused the cognitive impairment in a subset of these individuals.<sup>60</sup> Here, MCI was diagnosed using a neuropsychological battery and a clinical consensus. Being another memory test, it is perhaps unsurprising that the MST predicted MCI. However, the MST was not part of the clinical diagnosis. Therefore, its accuracy could aid primary care clinicians who only have access to tests like the MMSE. We hypothesize that including the MST in primary care could improve cognitive impairment identification and reduce MCI underdiagnosis. This is, as yet, untested, and it is not known whether the MST can differentiate subtypes of MCI.

#### 4.2 | Cognitive modeling of MST predicts AD biomarker status in cognitively healthy older adults

Recent studies utilized comprehensive cognitive batteries to identify individuals at higher risk of AD,<sup>14,30,61</sup> with longitudinal cognitive

testing used to identify healthy adults with elevated A $\beta$  and pTau levels.<sup>61-63</sup> Several use a shortened version of the MST (aka BPSO)<sup>61,62</sup> and longitudinal changes on this could better predict memory impairment over 3 months compared to baseline neuropsychological scores. Importantly, none of the other tasks within the cognitive battery could exceed baseline neuropsychological scores.

Early detection of AD-related pathology is paramount, but current clinical assessments are lagging pathologies, such as cerebral A $\beta$  deposition by up to 20 years, highlighting the need for more sensitive tasks.<sup>1,4</sup> A recent study demonstrated that combining multiple versions of the MST could modestly predict amyloid status in CN older adults.<sup>32</sup> Similarly, we found that MST performance modestly predicted A $\beta$  status (AUC = 0.64), and the MPT model enhanced this prediction to an AUC of 0.73. Our results affirm that MST performance relates to A $\beta$  status and show that inferred cognitive mechanisms improve its predictive value.

Amyloid accumulation, in the absence of pTau, does not correlate with cognitive impairment,<sup>64</sup> but pTau is associated with hippocampal hyperactivity, neurodegeneration, and the transition to dementia.<sup>6,65</sup> However, it is not known whether increased pTau levels can be predicted from performance on the MST. We found that, while the traditional MST metrics could not predict elevated pTau levels, integrating the MPT model improves prediction to an AUC of 0.71. This suggests that the MPT model can distinguish individuals with elevated pTau, highlighting its potential in identifying early markers of cognitive impairment associated with AD pathology.

#### 4.3 | Changes in guessing strategies as a function of cognitive impairment and biomarker status

When applying the MPT model to predict AD biomarkers and cognitive status, one emerging theme was the role of guessing strategies. Guessing strategies, or response biases, change in amnesic patients and individuals with dementia, so it is worthwhile to investigate how they change in people at risk of AD. Notably, individuals with increased biomarker levels tended to shift their guessing bias from "similar" to "old," and this was more pronounced among those with MCI, suggesting a continuum of change. This shift toward guessing "old" aligns with other work demonstrating that MCI is associated with more liberal response biases on REC tasks.<sup>66,67</sup> One plausible explanation is that individuals build up gist representations during a task and, unlike younger adults, cannot rely on item-level detail memory. Thus, without intact memory for details, individuals with MCI may overly rely on gist and show a more liberal response bias.<sup>67,68</sup> On the MST, this would cause the observed bias toward responding "old." One potential explanation for this shift is reduced dentate gyrus-based pattern separation, which causes a shift toward pattern completion, thereby increasing false memories.<sup>69,70</sup> Notably, these changes were more reliably seen using the MPT model than with raw traditional metrics. These outcomes underscore the superiority of cognitive modeling in recognizing individuals at risk of AD, supporting its utility as a digital biomarker for early detection of the disease.

The MST's three-choice nature precludes the use of cognitive models like signal detection theory or dual process models, which assume unidimensionality in latent memory strength. In these models, responses are treated as arising from memory strengths between "new" and "old" (too strong to respond "new" but not enough for "old"). People, however, can actively generate evidence for "similar" responses (e.g., recognizing a different image of the same object).<sup>23</sup> Despite this, work using these approaches on two-choice tasks has also found a shift toward a bias toward responding old in MCI and AD,<sup>71-73</sup> further suggesting that individuals at risk for AD exhibit changes in response patterns. There is, to our knowledge, one other cognitive model that has been developed for the three-choice MST<sup>74</sup> which uses linear ballistic accumulators to avoid the unidimensionality assumption of the MST.

Interestingly, different measures were important for predicting AD biomarkers compared to age or memory ability. This supports the idea that the mechanisms that contribute to pathological cognitive decline (ie, dementia) are not the same as the mechanisms that explain variability in memory ability in healthy older adults. The findings that LDI was the best predictor of age group and gist-based information was the best at separating NMD and AMD suggest that these measures are most related to memory ability in older adults, but not necessarily related to AD. Conversely, the shift in structured guessing (ie, response bias) may be more directly linked to the pathological processes in AD.

#### 4.4 | Limitations and future directions

This study is not without limitations, and several future directions could be taken to better assess the utility of cognitive modeling of the MST. One critical aspect to consider is how well cognitive modeling performs relative to other, standard neuropsychological scores. In prior work using these same data,<sup>50</sup> multiple neuropsychological episodic memory assessments showed no ability to discriminate amyloid of pTau181 status. However, this was not a clinical setting, and this direct comparison was not made.

Future work is also needed to determine how well MST performance measures could be used to create risk scores in the clinical setting. Ideally, this, combined with other measures, such as plasma, could help screen high-risk individuals and identify who may need more invasive and costly measures. Future research should address this gap and explore how the MST can be utilized in clinical settings.

## 5 | CONCLUSION

Here, we asked whether a MPT cognitive model of the MST could aid in the identification of individuals at risk of AD. We demonstrated that, in addition to predicting memory deficits and MCI, the MPT model of the MST could predict both amyloid and pTau status with AUCs of greater than 0.7 in older adults without signs of cognitive decline. This suggests that cognitive modeling of the MST holds significant potential as a non-invasive, efficient screening tool within the clinical setting.

## ACKNOWLEDGMENTS

We thank Alexandra Trelle, Anthony Wagner, and Elizabeth Mormino for graciously sharing their behavioral and CSF data from the Stanford Aging and Memory Study. This research was funded in part by R01 AG066683 (CS and ML) and P30 AG066519 (CS).

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest. Author disclosures are available in the [Supporting Information](#).

## CONSENT STATEMENT

All human subjects provided informed consent.

## ORCID

Casey R. Vanderlip  <https://orcid.org/0000-0002-8233-5829>

Craig E. L. Stark  <https://orcid.org/0000-0002-9334-8502>

## REFERENCES

1. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):280-292. doi:10.1016/j.jalz.2011.03.003
2. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14(4):535-562. doi:10.1016/j.jalz.2018.02.018
3. Jia J, Ning Y, Chen M, et al. Biomarker changes during 20 years preceding Alzheimer's disease. *N Engl J Med*. 2024;390(8):712-722. doi:10.1056/NEJMoa2310168
4. Li Y, Yen D, Hendrix RD, et al. Timing of biomarker changes in sporadic Alzheimer's disease in estimated years from symptom onset. *Ann Neurol*. 2024;95(5):951-965. doi:10.1002/ana.26891
5. Donohue MC, Sperling RA, Petersen R, et al. Association between elevated brain amyloid and subsequent cognitive decline among cognitively normal persons. *JAMA*. 2017;317(22):2305. doi:10.1001/jama.2017.6669
6. Ossenkoppele R, Pichet Binette A, Groot C, et al. Amyloid and tau PET-positive cognitively unimpaired individuals are at high risk for future cognitive decline. *Nat Med*. 2022;28(11):2381-2387. doi:10.1038/s41591-022-02049-x
7. Sperling RA, Karlawish J, Johnson KA. Preclinical Alzheimer disease—the challenges ahead. *Nat Rev Neurol*. 2013;9(1):54-58. doi:10.1038/nrneurol.2012.241
8. Boxer AL, Sperling R. Accelerating Alzheimer's therapeutic development: the past and future of clinical trials. *Cell*. 2023;186(22):4757-4772. doi:10.1016/j.cell.2023.09.023
9. McMahon PM, Araki SS, Sandberg EA, Neumann PJ, Gazelle GS. Cost-effectiveness of PET in the diagnosis of Alzheimer disease. *Radiology*. 2003;228(2):515-522. doi:10.1148/radiol.2282020915
10. Wittenberg R, Knapp M, Hu B, et al. The costs of dementia in England. *Int J Geriatr Psychiatry*. 2019;34(7):1095-1103. doi:10.1002/gps.5113
11. Hansson O, Blennow K, Zetterberg H, Dage J. Blood biomarkers for Alzheimer's disease in clinical practice and trials. *Nat Aging*. 2023;3(5):506-519. doi:10.1038/s43587-023-00403-3
12. Barthélemy NR, Salvadó G, Schindler S, et al. Highly accurate blood test for Alzheimer's disease comparable or superior to clinical CSF tests. *Nat Med*. 2024;30(4):1085-1095. doi:10.1038/s41591-024-02869-z
13. Dagum P. Digital biomarkers of cognitive function. *NPJ Digit Med*. 2018;1(1):10. doi:10.1038/s41746-018-0018-4

14. Macdougall A, Whitfield T, Needham K, Schott JM, Frost C, Walker Z. Predicting progression to Alzheimer's disease dementia using cognitive measures. *Int J Geriatr Psychiatry*. 2024;39(2):e6067. doi:10.1002/gps.6067
15. Ding Z, Lee TL, Chan AS. Digital cognitive biomarker for mild cognitive impairments and dementia: a systematic review. *J Clin Med*. 2022;11(14):4191. doi:10.3390/jcm11144191
16. Vanderlip CR, Stark CEL, Alzheimer's Disease Neuroimaging Initiative. Digital cognitive assessments as low-burden markers for predicting future cognitive decline and tau accumulation across the Alzheimer's spectrum. *bioRxiv*. 2024. Published online May 26, 2024. doi:10.1101/2024.05.23.595638
17. Wang W, Peng J, Hou J, et al. Predicting mild cognitive impairment progression to Alzheimer's disease based on machine learning analysis of cortical morphological features. *Aging Clin Exp Res*. 2023;35(8):1721-1730. doi:10.1007/s40520-023-02456-1
18. Hedden T, Oh H, Younger AP, Patel TA. Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology*. 2013;80(14):1341-1348. doi:10.1212/WNL.0b013e31828ab35d
19. Kirwan CB, Stark CEL. Overcoming interference: an fMRI investigation of pattern separation in the medial temporal lobe. *Learn Mem*. 2007;14:625-633.
20. Bakker A, Kirwan CB, Miller NI, Stark CEL. Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science*. 2008;319:1640-1642.
21. Lacy JW, Yassa MA, Stark SM, Muftuler LT, Stark CE. Distinct pattern separation related transfer functions in human CA3/dentate and CA1 revealed using high-resolution fMRI and variable mnemonic similarity. *Learn Mem*. 2011;18:15-18. doi:10.1101/lm.1971111
22. Brock Kirwan C, Hartshorn A, Stark SM, Goodrich-Hunsaker NJ, Hopkins RO, Stark CE. Pattern separation deficits following damage to the hippocampus. *Neuropsychologia*. 2012;50:2408-2414. doi:10.1016/j.neuropsychologia.2012.06.011
23. Stark SM, Kirwan CB, Stark CEL. Mnemonic Similarity Task: a tool for assessing hippocampal integrity. *Trends Cogn Sci*. 2019;23(11):938-951. doi:10.1016/j.tics.2019.08.003
24. Small SA, Perera GM, DeLaPaz R, Mayeux R, Stern Y. Differential regional dysfunction of the hippocampal formation among elderly with memory decline and Alzheimer's disease. *Ann Neurol*. 1999;45(4):466-472.
25. Small SA, Schobel SA, Buxton RB, Witter MP, Barnes CA. A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nat Rev Neurosci*. 2011;12(10):585-601. doi:10.1038/nrn3085
26. Morrison JH, Hof PR. Chapter 37. Selective vulnerability of corticocortical and hippocampal circuits in aging and Alzheimer's disease. In: *Progress in Brain Research*. Elsevier; 2002:467-486. doi:10.1016/S0079-6123(02)36039-4
27. Sabuncu MR. The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Arch Neurol*. 2011;68(8):1040-1048. doi:10.1001/archneurol.2011.167
28. Ally BA, Hussey EP, Ko PC, Molitor RJ. Pattern separation and pattern completion in Alzheimer's disease: evidence of rapid forgetting in amnesic mild cognitive impairment. *Hippocampus*. 2013;23(12):1246-1258. doi:10.1002/hipo.22162
29. Stark SM, Yassa MA, Lacy JW, Stark CE. A task to assess behavioral pattern separation (BPS) in humans: data from healthy aging and mild cognitive impairment. *Neuropsychologia*. 2013;51:2442-2449. doi:10.1016/j.neuropsychologia.2012.12.014
30. Papp KV, Rentz DM, Maruff P, et al. The computerized cognitive composite (C3) in A4, an Alzheimer's disease secondary prevention trial. *J Prev Alzheimers Dis*. 2021;8(1):59-67. doi:10.14283/jpad.2020.38
31. Belliard-Guérin G, Planche V. Mnemonic discrimination performance in a memory clinic: a pilot study. *J Alzheimers Dis*. 2023;94(4):1527-1534. doi:10.3233/JAD-230221
32. Kim S, Adams JN, Chappel-Farley MG, et al. Examining the diagnostic value of the mnemonic discrimination task for classification of cognitive status and amyloid-beta burden. *Neuropsychologia*. 2023;191:108727. doi:10.1016/j.neuropsychologia.2023.108727
33. Mohs R, Bakker A, Rosenzweig-Lipson S, et al. The HOPE4MCI study: a randomized double-blind assessment of AGB101 for the treatment of MCI due to AD. *Alzheimers Dement*. 2024;10(1):e12446. doi:10.1002/trc2.12446
34. Norman KA, Detre G, Polyn SM. Computational models of episodic memory. In: Sun R, ed. *The Cambridge Handbook of Computational Psychology*. 1st ed. Cambridge University Press; 2001:189-225. doi:10.1017/CBO9780511816772.011
35. Lee MD, Bock JR, Cushman I, Shankle WR. An application of multinomial processing tree models and Bayesian methods to understanding memory impairment. *J Math Psychol*. 2020;95:102328. doi:10.1016/j.jmp.2020.102328
36. Chwiesko C, Janecek J, Doering S, et al. Parsing memory and nonmemory contributions to age-related declines in mnemonic discrimination performance: a hierarchical Bayesian diffusion decision modeling approach. *Learn Mem*. 2023;30(11):296-309. doi:10.1101/lm.053838.123
37. Mulhauser K, Giordani B, Kavcic V, et al. Utility of diffusion modeling of cogstate brief battery test performance in detecting mild cognitive impairment. *Assessment*. 2023;30(3):847-855. doi:10.1177/107319112111069089
38. Lee MD, Stark CEL. Bayesian modeling of the Mnemonic Similarity Task using multinomial processing trees. *Behaviormetrika*. 2023;50(2):517-539. doi:10.1007/s41237-023-00193-3
39. Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA*. 1993;269(18):2386-2391.
40. Wechsler D. *Wechsler Memory Scale—Third Edition (WMS-III)*. The Psychological Corporation; 1997.
41. Rey A. L'examen psychologique dans les cas d'encephalopathie traumatique. (Les problems). The psychological examination in cases of traumatic encephalopathy. *Problems. Arch Psychol*. 1941;28:286-340. [In English]
42. Tombaugh T, Kozak J, Rees L. Normative Data Stratified by Age and Education for Two Measures of Verbal Fluency FAS and Animal Naming. *Arch Clin Neuropsychol*. 1999;14(2):167-177. doi:10.1016/s0887-6177(97)00095-4
43. Wechsler D. *Wechsler Adult Intelligence Scale (WAIS-III): Administration and Scoring Manual*. The Psychological Corporation; 1997.
44. Tombaugh TN. Trail making test A and B: normative data stratified by age and education. *Arch Clin Neuropsychol*. 2004;19(2):203-214. doi:10.1016/S0887-6177(03)00039-8
45. Rogalski EJ, Gefen T, Shi J, et al. Youthful memory capacity in old brains: anatomic and genetic clues from the northwestern SuperAging project. *J Cogn Neurosci*. 2013;25(1):29-36. doi:10.1162/jocn\_a\_00300
46. Radhakrishnan H, Bennett IJ, Stark CE. Higher-order multi-shell diffusion measures complement tensor metrics and volume in gray matter when predicting age and cognition. *NeuroImage*. 2022;253:119063. doi:10.1016/j.neuroimage.2022.119063
47. Harrison TM, Weintraub S, Mesulam MM, Rogalski E. Superior memory and higher cortical volumes in unusually successful cognitive aging. *J Int Neuropsychol Soc*. 2012;18(06):1081-1085. doi:10.1017/S1355617712000847
48. Gefen T, Shaw E, Whitney K, et al. Longitudinal neuropsychological performance of cognitive SuperAgers. *J Am Geriatr Soc*. 2014;62(8):1598-1600. doi:10.1111/jgs.12967
49. Gefen T, Peterson M, Papastefan ST, et al. Morphometric and histologic substrates of cingulate integrity in elders with exceptional memory capacity. *J Neurosci*. 2015;35(4):1781-1791. doi:10.1523/JNEUROSCI.2998-14.2015

50. Trelle AN, Carr VA, Wilson EN, et al. Association of CSF biomarkers with hippocampal-dependent memory in preclinical Alzheimer disease. *Neurology*. 2021;96(10):e1470-e1481. doi:10.1212/WNL.00000000000011477
51. Plummer MJAGS. A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), Vienna. 2003. 124, 1-10.
52. Seabold S, Perktold J. Published online 2010:5.
53. Raftery AE. Bayesian model selection in social research. *Sociol Methodol*. 1995;25:111-163. doi:10.2307/271063
54. Student. The probable error of a mean. *Biometrika*. 1908;6(1):1-25. doi:10.2307/2331554
55. Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. *Psychon Bull Rev*. 2004;11(1):192-196. doi:10.3758/BF03206482
56. Yassa MA, Stark SM, Bakker A, Albert MS, Gallagher M, Stark CEL. High-resolution structural and functional MRI of hippocampal CA3 and dentate gyrus in patients with amnesic mild cognitive impairment. *NeuroImage*. 2010;51(3):1242-1252.
57. Gellersen HM, Trelle AN, Henson RN, Simons JS. Executive function and high ambiguity perceptual discrimination contribute to individual differences in mnemonic discrimination in older adults. *Cognition*. 2021;209:104556. doi:10.1016/j.cognition.2020.104556
58. Mattke S, Jun H, Chen E, Liu Y, Becker A, Wallick C. Expected and diagnosed rates of mild cognitive impairment and dementia in the U.S. Medicare population: observational analysis. *Alzheimers Res Ther*. 2023;15(1):128. doi:10.1186/s13195-023-01272-z
59. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol*. 1999;56(3):303-308. doi:10.1001/archneur.56.3.303
60. Solfrizzi V, Panza F, Colacicco AM, et al. Vascular risk factors, incidence of MCI, and rates of progression to dementia. *Neurology*. 2004;63(10):1882-1891. doi:10.1212/01.WNL.0000144281.38555.E3
61. Lim YY, Snyder PJ, Pietrzak RH, et al. Sensitivity of composite scores to amyloid burden in preclinical Alzheimer's disease: introducing the z-scores of attention, verbal fluency, and episodic memory for nondemented older adults composite score. *Alzheimers Dement*. 2016;2(1):19-26. doi:10.1016/j.dadm.2015.11.003
62. Jutten RJ, Rentz DM, Fu JF, et al. Monthly at-home computerized cognitive testing to detect diminished practice effects in preclinical Alzheimer's disease. *Front Aging Neurosci*. 2022;13:800126. doi:10.3389/fnagi.2021.800126
63. Papp KV, Jutten RJ, Soberanes D, et al. Early detection of amyloid-related changes in memory among cognitively unimpaired older adults with daily digital testing. *Ann Neurol*. 2024;95(3):507-517. doi:10.1002/ana.26833
64. Desikan RS, McEvoy LK, Thompson WK, et al. Amyloid- $\beta$ -associated clinical decline occurs only in the presence of elevated P-tau. *Arch Neurol*. 2012;69(6):709-713. doi:10.1001/archneurol.2011.3354
65. Berron D, Cardenas-Blanco A, Bittner D, et al. Higher CSF tau levels are related to hippocampal hyperactivity and object mnemonic discrimination in older adults. *J Neurosci*. 2019;39(44):8788-8797. doi:10.1523/JNEUROSCI.1279-19.2019
66. Budson AE, Daffner KR, Desikan R, Schacter DL. When false recognition is unopposed by true recognition: gist-based memory distortion in Alzheimer's disease. *Neuropsychology*. 2000;14(2):277-287.
67. Budson AE, Desikan R, Daffner KR, Schacter DL. Perceptual false recognition in Alzheimer's disease. *Neuropsychology*. 2001;15(2):230-243.
68. Deason RG, Hussey EP, Ally BA, Budson AE. Changes in response bias with different study-test delays: evidence from young adults, older adults, and patients with Alzheimer's disease. *Neuropsychology*. 2012;26(1):119-126. doi:10.1037/a0026330
69. Leal SL, Yassa MA. Neurocognitive aging and the hippocampus across species. *Trends Neurosci*. 2015;38(12):800-812. doi:10.1016/j.tins.2015.10.003
70. Devitt AL, Schacter DL. False memories with age: neural and cognitive underpinnings. *Neuropsychologia*. 2016;91:346-359. doi:10.1016/j.neuropsychologia.2016.08.030
71. Snodgrass JG, Corwin J. Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J Exp Psychol Gen*. 1988;117(1):34-50. doi:10.1037/0096-3445.117.1.34
72. Bennett IJ, Golob EJ, Parker ES, Starr A. Memory evaluation in mild cognitive impairment using recall and recognition tests. *J Clin Exp Neuropsychol*. 2006;28(8):1408-1422. doi:10.1080/13803390500409583
73. Koen JD, Yonelinas AP. The effects of healthy aging, amnesic mild cognitive impairment, and Alzheimer's disease on recollection and familiarity: a meta-analytic review. *Neuropsychol Rev*. 2014;24(3):332-354. doi:10.1007/s11065-014-9266-5
74. Banavar NV, Noh SM, Wahlheim CN, et al. A response time model of the three-choice Mnemonic Similarity Task provides stable, mechanistically interpretable individual-difference measures. 2024;18. doi:10.31234/osf.io/yvdae

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Vanderlip CR, Lee MD, Stark CEL. Cognitive modeling of the Mnemonic Similarity Task as a digital biomarker for Alzheimer's disease. *Alzheimer's Dement*. 2024;20:6935-6947. <https://doi.org/10.1002/alz.14163>