medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Large-Scale Longitudinal Assessment of Declining Memory Function Through Individualized Modeling of Episodic Memory

Holly S. Hake[a,b], Maarten van der Velde[c,d,e], Bridget Leonard[a], Thomas Grabowski[f], Hedderik van Rijn[c,d,e], and *Andrea Stocco[a]

[a]Department of Psychology, University of Washington, Seattle, WA, USA
[b]Graduate Program in Neuroscience, University of Washington, Seattle, WA, USA
[c]SlimStampen B.V., Groningen, Netherlands
[d]Department of Experimental Psychology, University of Groningen, Groningen, Netherlands
[e]Behavioral and Cognitive Neurosciences, University of Groningen, Groningen, Netherlands
[f]Department of Radiology and Neurology, University of Washington, Seattle, WA, USA

Email: stocco@uw.edu

Author Contributions: Designed research: H.S.H., T.G., A.S.; Performed research: H.S.H., B.L., A.S.; Contributed new reagents/analytic tools: H.S.H., M.v.d.V., H.v.R., A.S.; Analyzed data: H.S.H., T.G., B.L., A.S.; Wrote the paper with contributions made from all authors: H.S.H., A.S.

Competing Interest Statement: The authors declare competing interests as follows: H.S.H., M.v.d.V., H.v.R., and A.S. are named on a preliminary patent related to the subject matter discussed in this article; M.v.d.V. and H.v.R. are employed by SlimStampen B.V., the company responsible for the development and commercialization of the AFLS (Adaptive Fact-Learning System).

Classification: Biological Sciences; Psychological and Cognitive Sciences

Keywords: memory; forgetting; dementia; diagnostics; ACT-R

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

The detection and tracking of progressive memory impairments, particularly in the context of neurodegenerative disorders, relies predominantly on traditional neuropsychological assessment and short cognitive screening tools. These methods, however, are resource-intensive and lack the accessibility and/or the repeatability necessary for effective early identification and tracking interventions. This study addresses the critical need for reliable and efficient diagnostic tools to track and predict memory decline in clinical settings. We demonstrate that an online, remote model-based memory assessment, can identify individuals with Mild Cognitive Impairment (MCI) with an accuracy rate exceeding 84% in a single 8-minute session. Furthermore, the test can be repeated multiple times with increasing accuracy over multiple assessments. The system's ability to monitor individual memory function inexpensively and longitudinally across various materials offers a robust and repeatable alternative to the static measures currently employed. Our findings show that traditional methods to assess memory decline could be replaced by adaptive, precise, and patient-friendly online tools based on computational modeling techniques. Moreover, our findings also open avenues for the proactive management of Alzheimer's disease and other dementias, as well as sensitively tracking the effect of interventions in early disease.

Significance Statement

We introduce, for the first time, a dependable, reliable, and repeatable model-based system for the online assessment of clinical memory impairment. Our findings demonstrate that this system can efficiently detect memory impairments using only 8 minutes of data collected online, marking a significant improvement over traditional assessments that typically require 3 hours in a clinical setting. These findings open up the possibility of inexpensive population-level monitoring of memory function.

Main Text

Introduction

Memory impairments, particularly prevalent in conditions such as dementia, pose a significant public health challenge. This challenge is exacerbated by aging populations and the high costs associated with care. The current diagnostic landscape for these impairments is fraught with limitations: tests are time-consuming, culturally biased, require specialized administration, and are vulnerable to practice effects. These factors collectively contribute to a significant underestimation of conditions such as Mild Cognitive Impairment (MCI), with studies revealing that only a small fraction of expected cases are accurately diagnosed in the United States. For example, one study estimated that only 8% of expected MCI cases are diagnosed (1), and similar studies confirm low detection rates (11-15%; (2, 3)), often attributing this to clinicians' limited skills and time constraints. This discrepancy highlights a critical gap in our healthcare system's ability to identify and address cognitive decline early on.

Despite these challenges, recent research indicates improvements in the detection rates of dementia, suggesting that these rates are now more closely aligned with expected figures (4). However, this progress is overshadowed by the continued oversight of the early stages of cognitive impairment. The early detection of such conditions is crucial, as many MCI causes are reversible, and lifestyle modifications can halt or reverse cognitive decline (5). The potential for early intervention to significantly alter patient outcomes underscores the pressing need for diagnostic tools that are not only more accurate but also accessible and patient-centric (6).

In response to these needs, an ideal diagnostic tool would bridge the existing gaps. It would provide reliable, interpretable data that accurately reflects an individual's memory function in real-world settings, distinguish between normal aging and pathological memory decline, facilitate frequent and unbiased testing, and require minimal intervention from specialists. Such a tool would be specifically designed to address memory-related concerns, be minimally influenced by external variables, and be easily usable by patients themselves.

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Moreover, the ideal diagnostic tool would be adaptable for remote use, aligning with the growing need for telemedicine. This feature would allow for the assessment of memory impairments from a distance, making it accessible to individuals who cannot easily visit healthcare facilities. Such remote administration capabilities would not only extend the tool's reach but also enhance patient convenience, safety, and inclusiveness.

Notably, traditional screening and diagnostic tools have not fully utilized contemporary computational frameworks of memory encoding and retrieval (7–9). In the present study, we pioneer using such a computational approach for diagnostic applications. Specifically, we utilize a well-established episodic memory model (7) in combination with an adaptive fact-learning system (AFLS) to create a unified diagnostic framework. The AFLS tailors stimulus presentation rates to individual users, and the computational model, iteratively calibrated through AFLS response data, evolves into a 'cognitive twin' of the user. This twin's internal parameters offer a quantified measure of the patient's latent memory capacity (10).

Central to our model is the *Speed of Forgetting* (SoF) parameter. The *SoF* quantifies the initial rate of episodic trace decay without rehearsal, setting the initial forgetting trajectory of a memory, and determining how much it could benefit from subsequent re-encodings and retrievals (7, 11, 12). Here, the *SoF* parameter is used to quantify an individual's memory function in real time. This aspect of our approach is a first in the field, enabling fast assessment of memory performance with unprecedented accuracy. Furthermore, by calibrating the *SoF* parameter using AFLS data, our method offers an individualized and continuously evolving assessment for each individual. This represents a significant advancement beyond the static, one-size-fits-all approach of traditional diagnostic tools.

Importantly, the *SoF* parameter's interpretation is transparent and straightforward, making it accessible for both clinicians and researchers. Its independence from standardization norms further enhances its utility, as it does not rely on conventional benchmarks that may not apply universally across diverse populations. This independence is crucial in ensuring that a diagnostic tool is broadly applicable and adaptable to various individual contexts. Thus, our model is not only aligned with the ideal criteria for diagnostic tools but also pushing forward the frontiers of understanding and assessing memory impairments.

To test this new approach, we conducted a longitudinal study of healthy elderly adults and elderly individuals with Mild Cognitive Impairment (MCI). We chose individuals with MCI, rather than dementia, to both ensure participants had sufficient cognitive abilities to perform the task and to test the model's ability to detect earlier, more subtle differences in memory function as it is often a precursor to AD and other forms of dementia (13). This cohort of individuals was followed for 6+ months, during which they performed weekly online model-based assessments to characterize their *Speed of Forgetting*. We hypothesized that (1) *SoF* values would be reliable across repeated assessments; (2) individuals with MCI would exhibit higher *SoF* values than healthy controls, with accelerated forgetting observed over time; (3) *SoF* values would have clinical validity, allowing for the identification of abnormal memory function; and finally, (4) *SoF* would primarily indicate consolidation processes rather than retrieval processes in memory formation.

Results

In our study, 24 individuals diagnosed with Mild Cognitive Impairment (MCI; 5 female) and 27 age-matched controls (19 female) participated in weekly assessments. These assessments were conducted through the AFLS website[1], allowing participants to engage from the comfort of their home on a computer, tablet, or smartphone (Fig. 1A). The assessments, termed "lessons" or "memory games" covered an array of topics such as Nature, History, Science, Art, Geography, Media, Food, and Culture, providing a diverse range of facts for memorization (Fig. 1B). The content was presented as cue-response pair associates, with responses provided through multiple-choice options or verbal recall. Throughout the study, participants completed up to 49 multiple-choice recognition sessions and four quarterly verbal recall sessions. The verbal recall

[1] adrc.memorylab.nl

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

sessions were conducted online using a video conferencing platform, with vocal responses automatically translated through Google's speech-to-text software, and subsequently checked by trained assistants.

***Speed of Forgetting* Shows Reliable Patterns Across Different Cognitive Materials.** Our findings demonstrated that the *Speed of Forgetting* (SoF) remained consistent over time, corroborating earlier observations by Sense et al. (14). Among the 51 subjects who completed >35 lessons, the *SoF* metrics showed a range from the lowest individual score of 0.23 to the highest individual score of 0.54 (Fig. 1C). This range, reflecting the full spread of *SoF* values across all participants and lessons, aligned with a normal distribution pattern (Fig. 1D). The lesson means for *SoF* ranged from 0.36-0.41 with a mean of 0.39 (SD = 0.0129). Although lesson averages showed limited variation from their respective means, there were slight differences in the *SoF* across various cognitive materials. This variability was more pronounced in some participants than others, implying that specific materials might be less memorable for certain individuals. Nevertheless, the reliability of the *SoF* parameter was confirmed by its test-retest reliability. Across all pairs of lessons, the *SoF* exhibited a Pearson correlation coefficient of r = 0.70, indicating consistent inter-lesson correlations (Fig. 1E-F). The reliability across different materials not only validates the *SoF* measure but also highlights its potential for repeated assessments in future studies.

***Speed of Forgetting* Reveals Clear Differences in Cognitive Performance Between Mild Cognitive Impairment and Healthy Individuals.** Next, we examined whether patients diagnosed with Mild Cognitive Impairment (MCI) would exhibit significantly higher *Speed of Forgetting* values than age-matched healthy controls. On average, MCI patients exhibited markedly higher *SoF* values than healthy controls, with healthy controls demonstrating a mean *SoF* of 0.37 +/- 0.02, notably lower than MCI patients' mean *SoF* of 0.42 +/- 0.02 (Welch's t = 47.65, $p < 0.0001$) (Fig. 2A-C).

Further analysis using a mixed linear model, which explained 56% of the variance, showed a significant main effect of diagnosis, with the MCI group displaying consistently higher *SoF* values ($\beta = 0.34$, $p < 0.001$). The time course of the study was examined by including both linear and quadratic terms for time in the model. While the linear progression of weeks did not significantly impact *SoF* ($\beta = -0.03$, 95% CI [-0.15, 0.08], $p = 0.588$), the quadratic term approached significance ($\beta = 0.08$, 95% CI [-0.02, 0.17], $p = 0.107$), indicating a potential nonlinear effect of time on *SoF*. Notably, the interaction between MCI diagnosis and the quadratic time term was significant ($\beta = -0.14$, 95% CI [-0.26, -0.03], $p = 0.015$), suggesting a stronger nonlinear pattern of memory decline over time in MCI patients (Fig. 2D).

In addition to the main effects, age ($\beta = 0.0016$, $p < 0.005$) and sex ($\beta = 0.020$, $p = 0.02$) were significant factors influencing *SoF*, with older age and male sex linked to an increased rate of forgetting. The interaction between MCI status and age was also significant ($\beta = -0.00$, 95% CI [-0.01, -0.00], $p < 0.001$), indicating that age modifies the relationship between MCI and *SoF*, albeit with a minimal effect size.

Finally, to explore our approach's ability to parse out the nuances in cognitive deficits, we categorized the MCI participants into distinct MCI subtypes based on their initial diagnosis: amnestic single domain (aMCI S), amnestic multiple domains (aMCI M), and non-amnestic single MCI (naMCI S). Both amnestic subtypes exhibit impairments in memory, whereas the naMCI subtype demonstrates intact memory functionality (15). Our results from a Kruskal-Wallis rank sum test revealed significant *SoF* discrepancies between subtypes ($\chi^2(4) = 480$, $p < 0.001$), with pairwise comparisons indicating lower *SoF*s in healthy controls versus all MCI subtypes and distinctive *SoF* values among MCI subtypes (Fig. 2E). These findings elucidate that naMCI S individuals demonstrate cognitive abilities that lie between those of healthy controls and MCI patients, consistent with their language processing impairments and preserved memory functions.

***Speed of Forgetting* as a High-Validity Diagnostic Tool for Cognitive Impairment.** To evaluate the *Speed of Forgetting* as a diagnostic tool for cognitive impairment, we first compared

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

it to the Montreal Cognitive Assessment (MoCA), the most commonly used screening questionnaire for MCI diagnosis (16). As expected, we found a pronounced negative correlation ($r = -0.71$, $p < 0.0001$) between MoCA scores and *SoF*, confirming the predictive validity of *SoF* as a clinically applicable cognitive performance metric (Fig. 3A).

Next, we quantified the probability of MCI diagnosis given a participant's *SoF* values by binning results from individual sessions, using bins ranging from 0.25 to 0.55 in increments of 0.025. For each bin, we computed the proportion of scores coming from MCI individuals. Our analysis revealed that the probability of an MCI diagnosis grows as a logistic function of *SoF*, with an inflection point at 0.40 (Fig. 3B, red dotted line). A logistic model fitted to the data accounted for 42% of the variance in diagnosis probability. Additionally, we identified an optimal *SoF* threshold value of 0.388 using Youden's J statistic. At this threshold, the model demonstrated a specificity of 0.747 and a sensitivity of 0.741, effectively distinguishing between MCI and healthy controls.

We further scrutinized the diagnostic precision of *SoF* by employing it as a binary classifier between healthy controls (HC) and MCI cases. Receiver Operating Characteristic (ROC) analysis incorporating *SoF* measurements from approximately 2000 individual sessions revealed that a single-session *SoF* can achieve a classification accuracy of 84% (Fig. 3C).

Comparatively, *SoF* outperformed isolated behavioral metrics derived from the AFLS, including accuracy and response time, emphasizing the added value of a model-based interpretation of data. It is important to mention that comparing *SoF* with accuracy and response time in this paradigm can pose challenges due to their common origin in the AFLS. This can lead to inflated accuracies and response times that could skew comparative analyses. Despite this, however, the superior performance of *SoF* underscores its utility as a dependable measure for detecting cognitive impairment.

Finally, we explored the impact of multiple assessments on classification accuracy. While participants' total mean *SoF* across all 49 recognition sessions indicated a higher diagnostic accuracy of 94% (Fig. 3D), ROC curves generated from just the first 10 sessions revealed peak predictive accuracy at four assessments (Fig. 3E). This key finding suggests that only four assessments are needed to achieve the same level of classification accuracy as all 49 recognition assessments.

Recall Tests Confirm *Speed of Forgetting* Captures Consolidation Rather than Retrieval Processes. In our final analysis, we examined the recall sessions in contrast with the recognition sessions to ascertain the underlying memory processes (Fig. 4A-C). We observed significantly lower *Speed of Forgetting* values in recall compared to recognition (paired $t(45) = -10.90$, $p < 0.0001$), aligning with the notion that, while free verbal recall may be more difficult, recall tests facilitate better encoding through deeper information processing (17).

Using a random slope mixed linear model, incorporating clinical diagnosis and test type as factors, we noted a significant interaction between test type and diagnosis ($p < 0.001$). Specifically, individuals with MCI showed diminished benefits from recalling items compared to healthy controls. Despite theoretical distinctions between recall and recognition memory (18), *SoF* values exhibited a strong correlation across all participants for both test types (Pearson $r(44) = 0.86$, $t = 11.34$, $p < 0.0001$), suggesting a common mechanism of memory consolidation processing, rather than retrieval mode specificity.

Further analysis of *SoF*'s diagnostic ability for recall sessions revealed comparable classification accuracy to recognition *SoF*, as indicated by ROC AUC values of 91-93% (Fig. 4D-F). The slightly higher AUCs in recall sessions likely stem from increased separation between MCI and healthy controls. This consistency between recall and recognition assessments underscores the robustness of *SoF* in capturing memory consolidation processes and its consistent diagnostic utility across different memory retrieval modes.

Discussion

This study introduces a model-based, online assessment as a novel approach for diagnosing and monitoring memory impairments, such as Mild Cognitive Impairment (MCI). By implementing a computational model that dynamically adjusts to individual forgetting patterns—captured through

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

the novel metric, the *Speed of Forgetting* (SoF)—we have developed a more personalized solution for early memory impairment detection. Our methodology not only addresses various limitations of traditional neuropsychological assessments but also offers enhanced efficiency and scalability, particularly in handling large-scale data collection, making it a versatile tool for seamless widespread implementation.

***Speed of Forgetting* as a Novel Memory Metric.** The *SoF*, central to our findings, emerges as a robust metric for memory decline, validated across a variety of cognitive materials and memory assessments. Its high test-retest reliability substantiates its role as a reliable measure of memory function, while its clarity and interpretability render it a valuable asset for both clinical and research applications.

The study design, spanning weekly assessments over a six-to-twelve-month period, provided unique insights into the trajectory of cognitive decline in MCI, crucial for identifying and intervening in early-stage neurodegenerative conditions. The statistical analysis of MCI interaction with time illustrated a complex, nonlinear relationship ($\beta = -0.14$, $p = 0.015$), suggesting a decline in memory with patterns that change over time, which could be pivotal for tailoring individual treatment plans (Fig. 2D).

Our study also offers a granular look at MCI subtypes. The differential cognitive profiles within MCI subgroups, especially highlighted in non-amnestic MCI individuals with language processing intact, suggest a more intricate understanding of MCI (Fig. 2E). Such specificity in diagnosis is particularly relevant given the heterogeneity of neurodegenerative diseases like Alzheimer's and the need for assessments that transcend traditional reliance on verbal memory metrics.

The *SoF*'s diagnostic accuracy—84% for single sessions and 94% for multiple sessions—strongly advocates for its adoption into clinical practice (Fig 3C, D). Moreover, *SoF*'s design for remote administration is timely, paralleling the rise of telehealth solutions and meeting the demand for accessible cognitive assessment tools in the global pivot towards digital health services for patient care and monitoring.

Comparison to Other Cognitive Assessment. When comparing our *SoF*-based method with the Montreal Cognitive Assessment (MoCA; (16)), *SoF* not only matches the accuracy of MoCA but also advances the field in terms of test administration and frequency (Fig. 3A). This is particularly notable due to the potential risks associated with using brief cognitive screening tools, such as overdiagnosis and underdiagnosis, especially when relying solely on a single diagnostic test (19, 20). Moreover, repeated measuring using *SoF* offers a unique advantage, particularly in tracking cognitive fluctuations, such as "lucid days," which may not be effectively captured by single assessments. These fluctuations provide valuable insights into the dynamic nature of cognitive functioning and aid clinicians and researchers in understanding the progression of cognitive disorders over time. Additionally, while brief assessments like the MoCA excel in identifying substantial cognitive impairments, they may overlook early or subtle deficits and struggle to determine their underlying causes. Furthermore, comprehensive neuropsychological evaluations, although considered the gold standard, are often unavailable in community-based or primary care settings.

Underlying Memory Processes. A crucial evaluation of our model revolved around whether the *SoF* parameter effectively tracks error rates or effective forgetting. In one scenario, we would anticipate a higher *SoF* for the recall condition (due to increased error rates), while in the other, we would expect it to be lower (attributable to enhanced encoding). To validate this prediction, we included a recall version of the assessment where recognition options were omitted, and responses were recorded via a voice-activated microphone.

In our investigation into recall versus recognition memory processes, we observed lower *SoF*s in recall conditions (Fig. 4C), reinforcing the idea that recall tasks, while more challenging, enhance memory encoding and retention (17). These findings are consistent with established literature and affirm the memory processes in our model, as evidenced by the substantial correlation between *SoF* values from recall and recognition trials (Pearson $r = 0.85$, $p < 0.001$).

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

This strong correlation is also beneficial because the recall versions are more difficult to administer remotely

Limitations and Considerations. While the study demonstrates the effectiveness of the adaptive fact-learning system in detecting memory impairments, it also revealed certain limitations.

First, the *Speed of Forgetting,* in essence, measures what we believe to be the speed of "passive" forgetting. Passive forgetting, i.e. the loss of information over time due to the passage of time rather than a deliberate attempt to forget, could be due to loss of context clues, retrieval interference from other similar memories, and "natural" biological decay (21). Critically, some of these processes are accelerated in aging (22) and abnormally elevated in amnestic dementias, such as AD. More research is needed to determine which of these passive mechanisms is most closely linked to this parameter.

However, it's important to note that a similar criticism can be made of other existing assessments. They all operate at the level of observable performance, so compared to them, our study takes a step further in the direction of mechanistic explanations of performance. By focusing on the 'forgetting' process, we delve into the underlying mechanisms contributing to memory decline, which provide valuable insights beyond mere performance metrics and could thus be used to examine existing debates in the field regarding the contributions of various biological processes.

Second, while the *SoF* model parameter is theoretically independent of the assessments used in this study and can be adapted to data from other memory assessment tools, it remains tethered to a specific model of episodic memory. This underlying computational model is derived from a Bayesian analysis of memory (23) combined with the Multiple Trace Theory (24) and is just one conceptual framework for understanding memory, not the sole model.

Third, while our model shows promise in assessing early memory impairment, it has not been tested on participants with dementia. This is a significant limitation, as dementia presents unique challenges in cognitive functioning that may not be adequately captured by the model's current capabilities. Additionally, like any test, there are limits to the applicability of our model across diverse patient populations. Patients must still be able to perform certain basic tasks for the model to effectively assess memory impairment. Factors such as motor skills, language comprehension, and attention span may impact the utility of the model in accurately assessing cognitive decline.

However, it's worth mentioning that in the analysis of language-impaired non-amnestic MCI patients, *SoF* emerged as a more precise measure of memory impairment, suggesting it can still provide valuable insights for certain subsets of conditions.

Overall, further refinement and validation of the model in diverse populations is crucial for enhancing its clinical utility and reliability.

Future Directions and Applications. Despite these limitations, the *Speed of Forgetting* emerges as a reliable and highly valid metric for diagnosing cognitive impairment, surpassing traditional assessment tools in utility due to its online, unsupervised nature and adaptability for frequent use. These findings hold promising implications for the field of computational psychiatry and for the broader community. This paradigm can facilitate the collection of reliable data from populations that are often underrepresented in research and lack easy access to professional clinical assessment, including ethnic, racial, and linguistic minorities, as well as individuals residing in rural, low-income, or remote regions.

The remarkable repeatability and stability of the *SoF* model parameter position it as an exceptional tool for evaluating the effectiveness of interventions such as newly FDA-approved immunotherapies (25), and other neuromodulation tools and cognitive enhancers. Furthermore, by integrating the *SoF* model with neuroimaging data, it is possible to uncover the underlying neural mechanisms responsible for memory decline in aging and disease. Recent studies by Zhou et al. (26) and Xu et al. (27) have demonstrated this approach utilizing previous iterations of

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

the SoF task to analyze differences in functional connectivity networks and pinpoint specific brain regions correlated with forgetting.

Lastly, the *SoF*'s ability to track memory changes over time is particularly valuable for the early detection of MCI, crucial in delaying the onset of Alzheimer's Disease and related conditions. These combined attributes establish the *SoF* memory metric as a potent tool in both research and clinical practice.

Materials and Methods

Participants. Fifty-one participants were recruited on a rolling basis from the University of Washington Alzheimer's Disease Research Center. The inclusion criteria for the study were as follows: (1) age between 55 and 85 years, (2) fluency in English, and (3) no major medical or psychiatric conditions that would affect cognitive performance. Participants were classified into two groups: healthy cognition (N = 27; 19 female aged 58-84, 5 male aged 60-84) and those with Mild Cognitive Impairment (MCI; N = 24; 5 female aged 64-77, 19 male aged 68-83). All participants provided informed consent and were compensated for their participation in the online memory game portion of the study. All recruitment and testing procedures were approved by the University's Institutional Review Board.

Patient Selection and Diagnosis. Mild Cognitive Impairment (MCI) has commonly been used to describe the intermediary stage between normal cognitive aging and dementia (28). It can be defined as a decline in cognitive abilities that is greater than what is typical for a person's age and educational background (1−1.5 standard deviations below normative expectations) but does not meet the criteria for a diagnosis of dementia (29). MCI was diagnosed using a combination of methods including clinical evaluation, cognitive testing, and medical history. The clinical evaluation was conducted by a geriatric psychiatrist or a neurologist, who assessed the participant's cognitive and functional abilities using standardized tools. Cognitive testing was performed using a battery of neuropsychological tests that measured various cognitive domains such as memory, attention, and executive function. Medical history was obtained through a structured interview and review of medical records. Participants were classified as having MCI if they had a Clinical Dementia Rating scale <= 0.5. Additionally, individuals with subjective reports of decline by self and/or informant in conjunction with objective cognitive deficits were also included in the MCI group. Healthy controls were screened for cognitive impairment using the same methods as MCI participants. They were classified as healthy controls if they scored within normal limits on cognitive tests and had no history of cognitive decline or functional impairment.

All participants with memory impairment were classified as having amnestic Mild Cognitive Impairment (MCI) at the moment of enrollment. During the study, one participant was reclassified as a non-amnestic MCI (naMCI), one was diagnosed with Alzheimer's Disease (AD) upon re-assessment, and one was diagnosed as having Autism Spectrum Disorder (ASD) in addition to memory impairments. Of the 19 individuals who were classified as amnestic MCI, seven exhibited deficits only in the memory domain (aMCI Single), and 11 exhibited deficits in multiple domains (aMCI Multiple), such as executive function, language, or visual reasoning (15).

Episodic Memory Model. The model used herein was originally developed by Anderson & Schooler (7) and later expanded by Pavlik and Anderson (11) and van Rijn (12). Consistent with the Multiple Trace Theory (24), the model assumes that a memory is made of individual traces created every time the same information is encountered. Each trace decays according to the power law of forgetting (30). The odds of retrieving a memory *m* at time *t* are proportional to its activation *A*(*m*, *t*), which represents the log odds of retrieving any of its component traces, as shown in Eq. (1).

$$A(m, t) = \log \sum_i (t - t(i))^{-d(i)} \quad (1)$$

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Where $t(i)$ is the creation time of the i-th trace, and $d(i)$ is its characteristic power decay rate. This trace-specific decay rate, in turn, depends on the residual activation of the memory at the time the trace was created (11, 14):

$$d(i) = e^{A(m,\, t = t(i))} + \varphi \qquad (2)$$

Because Eq. (2) makes the decay rate of each trace dependent on the memory's activation, it provides an explanation for the spacing effect (31). Traces closer in time have higher decay rates because of the memory's greater activation A(m, t) at time t(i). Note that this model still depends on a single parameter, φ, which we refer to as the *Speed of Forgetting* (SoF). Thus, the odds of being able to recall a memory at a later time depend solely on the rate at which the memory is forgotten. Additionally, this suggests that by looking at the history of a memory and the number of times it has been assessed, it is possible to determine the rate at which that memory is forgotten.

Adaptive Memory Assessment. Weekly at-home assessments were completed with the online adaptive fact-learning system (AFLS) described in Sense et al. (14). This system continuously estimates the individualized *Speed of Forgetting* values in real time as the participant works through the lesson. The software was designed so that participants could perform the task from home using any mobile device. The AFLS works by presenting new study pairs (e.g., "France / Paris") and scheduling repeated tests (e.g., "France / ?") at strategic points based on the online estimates of a user's *Speed of Forgetting*. Figure 1A provides an example of the software interface. Test probes included only the cue of a pair, e.g. "France / ?"; participants responded by selecting one of four options on a screen.

Study Materials. Fifty-three lessons were prepared in advance, spanning different topics (such as European capitals, Swahili words, Asian flags, bird species, types of pasta, and flower species). The materials were vetted before the experiment to make sure they were comparable in terms of familiarity and difficulty. For each lesson, 15 different pairs were created, each of which associated an object with an English noun. In half of the pairs, the object was presented as an image (e.g., a picture of a pasta shape with the name "Fettuccine" for the Pasta lesson), and in the other half, the object was a word (e.g., "France" / "Paris" for European capitals). The number of pairs studied in each lesson depended on the response times and errors of the individual.

Data Processing. The repetition, activation, and *Speed of Forgetting* values for each term were calculated using functions from the AFLS. The average *Speed of Forgetting* values for each lesson and the individual were identified by using the terminal φ value of each pair at the very last repetition of that term. The data was then filtered to only contain the first full session of a topic (>6 min). This was needed to eliminate any superfluous sessions (some participants desired to complete the task more than once). The data was also organized by the week the lesson was completed to view temporal trends.

Statistics. Behavioral data were assessed for outliers defined as being outside of 2.5×SD of the mean and were subsequently excluded from analysis. Further statistical analysis was handled in R (v4.2.2) using the R Studio GUI (version 2022.7.2.576).

Data, Materials, and Software Availability. Our online repository containing both raw data files and analyzed data can be accessed at our online OSF Repository.

ACKNOWLEDGMENTS. We would like to thank Theresa Kehne and Lali Ghate from the Alzheimer's Disease Research Center (ADRC) for their help in recruiting the participants, as well as all participants and student helpers for their contributions. We thank Thomas Wilschut from SlimStampen B.V. for his support in the design of the recall sessions; and Alyssa Williams from Mississippi State University and Anais Capik, Sara Ulibarri, and Nevada Simpson from the

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

University of Washington for their support in collecting data from the recall sessions. This work was supported by a pilot award from the Garvey Institute for Brain Health Solutions.

References

1. Y. Liu, H. Jun, A. Becker, C. Wallick, S. Mattke, Detection Rates of Mild Cognitive Impairment in Primary Care for the United States Medicare Population. *J Prev Alzheimers Dis* **11**, 7–12 (2024).

2. L. White, *et al.*, Observational study of patient characteristics associated with a timely diagnosis of dementia and mild cognitive impairment without dementia. *J. Gen. Intern. Med.* **37**, 2957–2965 (2022).

3. G. M. Savva, A. Arthur, Who has undiagnosed dementia? A cross-sectional analysis of participants of the Aging, Demographics and Memory Study. *Age Ageing* **44**, 642–647 (2015).

4. Y. Zhu, Y. Chen, E. M. Crimmins, J. M. Zissimopoulos, Sex, Race, and Age Differences in Prevalence of Dementia in Medicare Claims and Survey Data. *J. Gerontol. B Psychol. Sci. Soc. Sci.* **76**, 596–606 (2021).

5. T. Ngandu, *et al.*, A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. *Lancet* **385**, 2255–2263 (2015).

6. J. Gallée, J. Cartwright, A. Volkmer, A. Whitworth, D. Hersh, “Please Don’t Assess Him to Destruction”: The R.A.I.S.E. Assessment Framework for Primary Progressive Aphasia. *Am. J. Speech. Lang. Pathol.* **32**, 391–410 (2023).

7. J. R. Anderson, L. J. Schooler, Reflections of the Environment in Memory. *Psychol. Sci.* **2**, 396–408 (1991).

8. D. L. Hintzman, Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychol. Rev.* **95**, 528–551 (1988).

9. R. M. Shiffrin, M. Steyvers, A model for recognition memory: REM—retrieving effectively from memory. *Psychon. Bull. Rev.* **4**, 145–166 (1997).

10. S. Somers, A. Oltramari, C. Lebiere, Cognitive twin: A cognitive approach to personalized assistants (2020) (January 27, 2024).

11. P. I. Pavlik Jr, J. R. Anderson, Practice and forgetting effects on vocabulary memory: an activation-based model of the spacing effect. *Cogn. Sci.* **29**, 559–586 (2005).

12. H. Van Rijn, L. van Maanen, M. van Woudenberg, Passing the test: Improving learning gains by balancing spacing and testing effects in *Proceedings of the 9th International Conference of Cognitive Modeling*, (2009), pp. 7–6.

13. R. C. Petersen, *et al.*, Mild cognitive impairment: ten years later. *Arch. Neurol.* **66**, 1447–1455 (2009).

14. F. Sense, F. Behrens, R. R. Meijer, H. van Rijn, An Individual’s Rate of Forgetting is Stable Over Time but Differs Across Materials. *Top. Cogn. Sci.* **8**, 305–321 (2016).

15. R. C. Petersen, *Mild Cognitive Impairment: Aging to Alzheimer’s Disease* (Oxford University Press, 2003).

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

16. Z. S. Nasreddine, *et al.*, The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **53**, 695–699 (2005).

17. M. Moscovitch, F. I. M. Craik, Depth of processing, retrieval cues, and uniqueness of encoding as factors in recall. *Journal of Verbal Learning and Verbal Behavior* **15**, 447–458 (1976).

18. C. Ranganath, K. A. Paller, Neural correlates of memory retrieval and evaluation. *Brain Res. Cogn. Brain Res.* **9**, 209–222 (2000).

19. H. Brodaty, *et al.*, Operationalizing the Diagnostic Criteria for Mild Cognitive Impairment: The Salience of Objective Measures in Predicting Incident Dementia. *Am. J. Geriatr. Psychiatry* **25**, 485–497 (2017).

20. J. M. Ranson, *et al.*, Predictors of dementia misclassification when using brief cognitive assessments. *Neurol. Clin. Pract.* **9**, 109–117 (2019).

21. R. L. Davis, Y. Zhong, The biology of forgetting—A perspective. *Neuron* **95**, 490–503 (2017).

22. Y. Shuai, *et al.*, Forgetting is regulated through Rac activity in Drosophila. *Cell* **140**, 579–589 (2010).

23. J. R. Anderson, *The adaptive character of thought* (Routledge, 1990).

24. L. Nadel, A. Samsonovich, L. Ryan, M. Moscovitch, Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* **10**, 352–368 (2000).

25. Lecanemab Approved for Treatment of Early Alzheimer's Disease. *Alzheimer's Disease and Dementia* (March 12, 2024).

26. P. Zhou, F. Sense, H. van Rijn, A. Stocco, Reflections of idiographic long-term memory characteristics in resting-state neuroimaging data. *Cognition* **212**, 104660 (2021).

27. Y. Xu, C. Prat, F. Sense, H. van Rijn, A. Stocco, Distributed Patterns of Functional Connectivity Underlie Individual Differences in Long-Term Memory Forgetting. *bioRxiv*, 2021.08.04.455133 (2021).

28. M. S. Albert, *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. *Alzheimers. Dement.* **7**, 270–279 (2011).

29. B. Winblad, *et al.*, Mild cognitive impairment--beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J. Intern. Med.* **256**, 240–246 (2004).

30. A. Newell, P. S. Rosenbloom, Mechanisms of skill acquisition and the law of practice (1990) https:/doi.org/10.1184/R1/6607196.V1.

31. N. J. Cepeda, E. Vul, D. Rohrer, J. T. Wixted, H. Pashler, Spacing effects in learning: a temporal ridgeline of optimal retention. *Psychol. Sci.* **19**, 1095–1102 (2008).

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figures and Tables

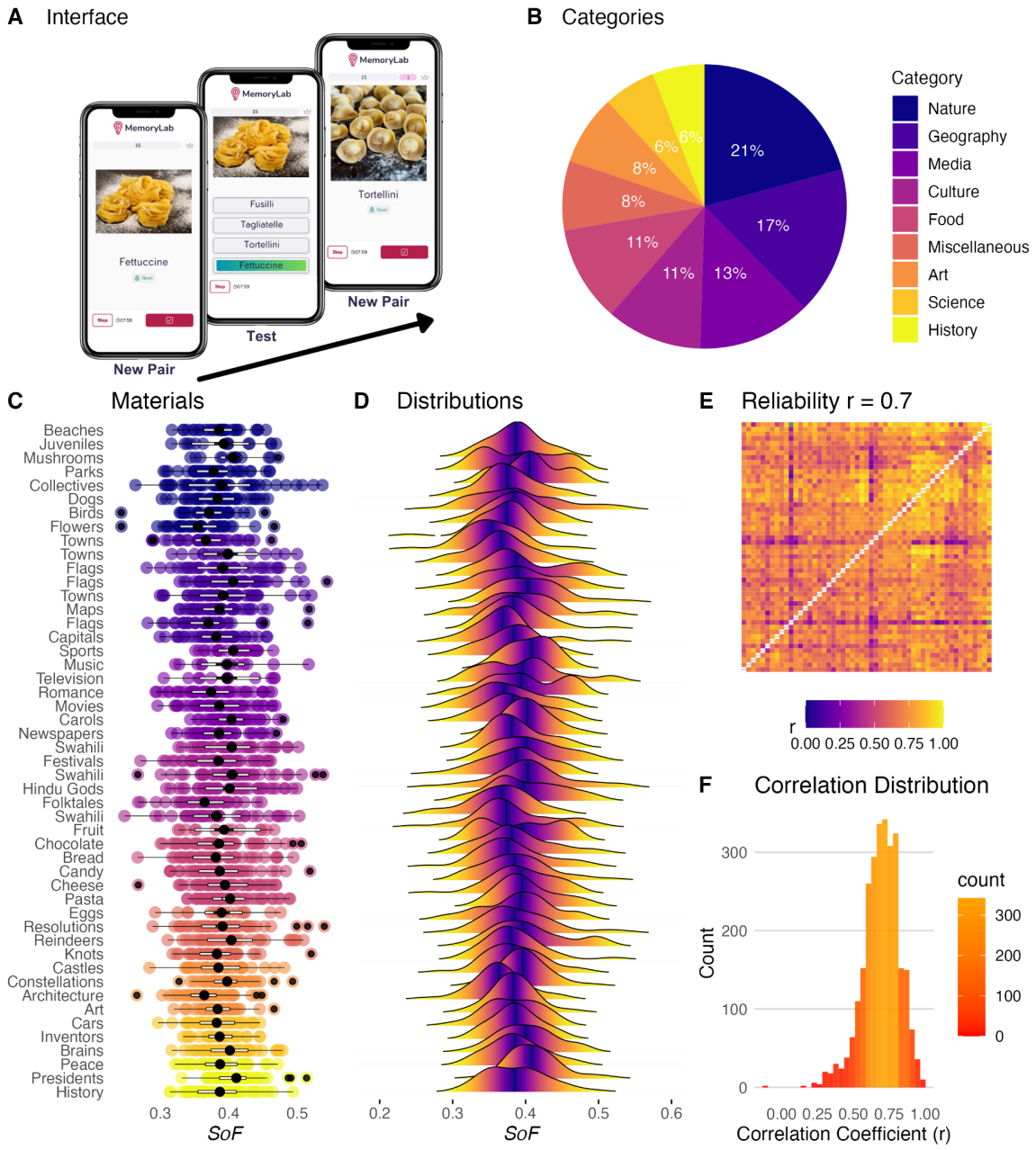

Figure 1. Multifaceted Analysis of *Speed of Forgetting* Across Diverse Categories.
(A) Mobile interface of the MemoryLab recognition task. (B) Categories. Pie chart shows the distribution of lessons into nine categories. (C) Materials. Scatterplot with the horizontal axis representing individual participant *Speeds of Forgetting* (SoF) and data points colored by category. (D) Distribution. Ridge plot illustrates the distribution of *SoF* across recognition lessons, with each line representing a lesson and color intensity reflecting tail probabilities. (E) Reliability of *SoF*. The heatmap represents the Pearson correlations between the individual *SoF* values of different lessons. The mean correlation coefficient (r) is 0.7. (F) Correlation Distribution. Histogram displaying the distribution of correlation coefficients from the heatmap in E.

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

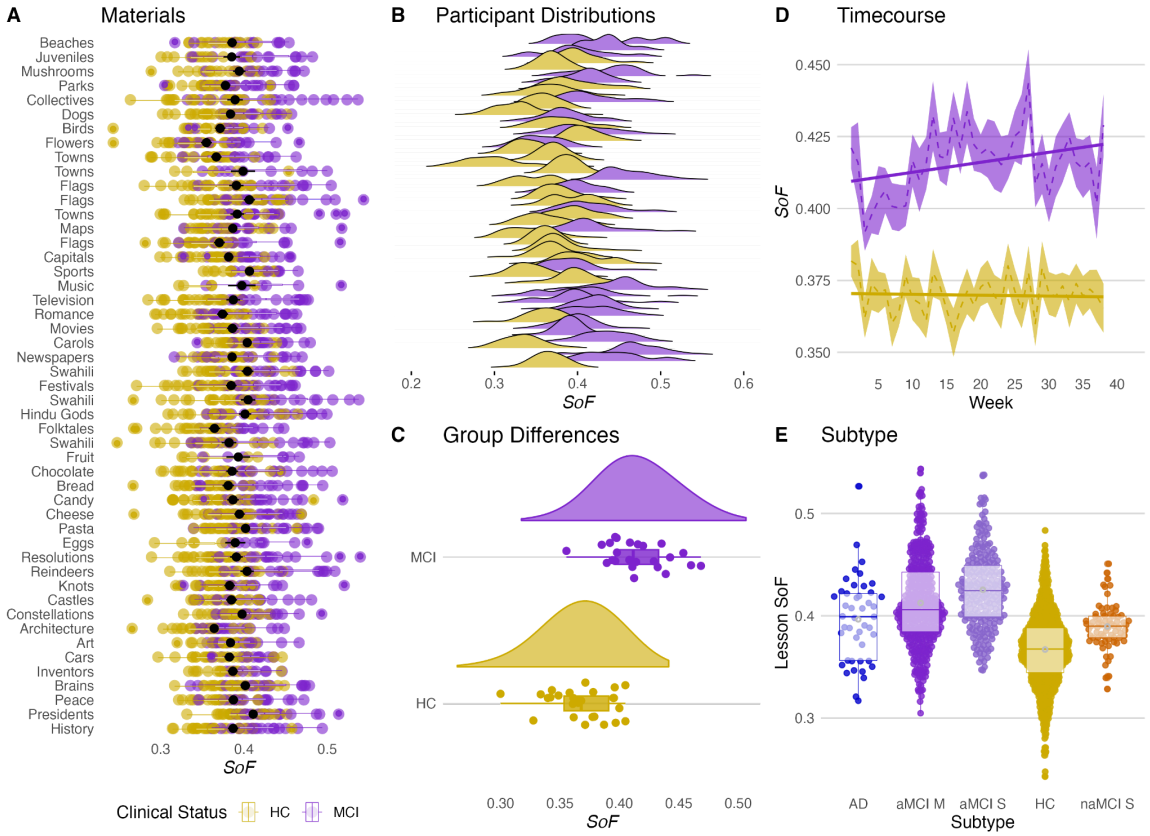

Figure 2. *Speed of Forgetting* Reveals Clear Differences in Cognitive Performance Between Mild Cognitive Impairment and Healthy Individuals.
(A) Materials. Scatterplot with the horizontal axis representing *Speed of Forgetting* (SoF), color-coded by clinical status: Healthy Control (HC) and Mild Cognitive Impairment (MCI). (B) Participant Distributions. Ridge plots illustrate the distribution of *SoF* among participants for all recognition lessons they completed. (C) Group Differences. Rain plots showing group differences in *SoF* between HC and MCI, indicating distinct memory retention capabilities between the groups. (D) Timecourse. Line graph displaying the time course of *SoF* over 40 weeks, with shaded areas representing 95% confidence intervals, for both HC (yellow) and MCI (purple) groups. (E) Subtype. Beeswarm plots for AD (dark blue), aMCI M (dark purple), aMCI S (light purple), HC (yellow), and naMCI (orange), with each point reflecting an individual lesson. Median and quartiles are summarized by boxplots, and data density by violin plots.

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

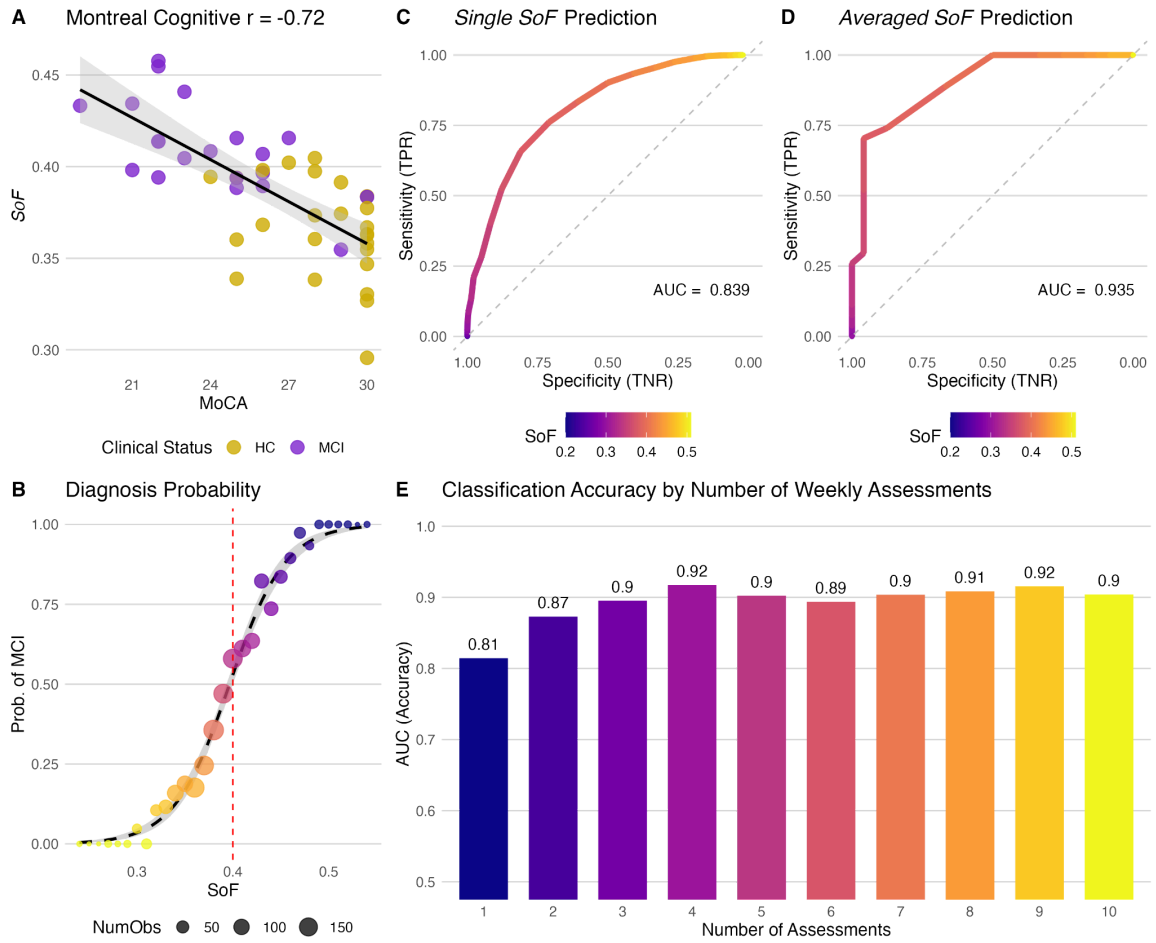

Figure 3. *Speed of Forgetting* as a High-Validity Diagnostic Tool for Cognitive Impairment. (A) Gold Standard. Scatter plot correlating *SoF* with Montreal Cognitive Assessment (MoCA) scores, showing a significant negative correlation (r = -0.71). (B) Diagnosis Probability. Sigmoid curve correlating *SoF* scores with the likelihood of MCI diagnosis, where point size indicates the number of observations contributing to each data point. (C) ROC curve for predicting cognitive status using a single *SoF* recognition measurement, with the area under the curve (AUC) demonstrating the model's accuracy, and individual points representing varying *SoF* thresholds. (D) Improved prediction model using the average *SoF*, as shown by higher AUC values, suggesting better performance with aggregated *SoF* data. (E) Multiple ROC curve AUC values represent the accuracy of predictions based on the number of recognition assessments, illustrating the incremental benefit of repeated measurements.

medRxiv preprint doi: https://doi.org/10.1101/2024.03.15.24304345; this version posted March 16, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

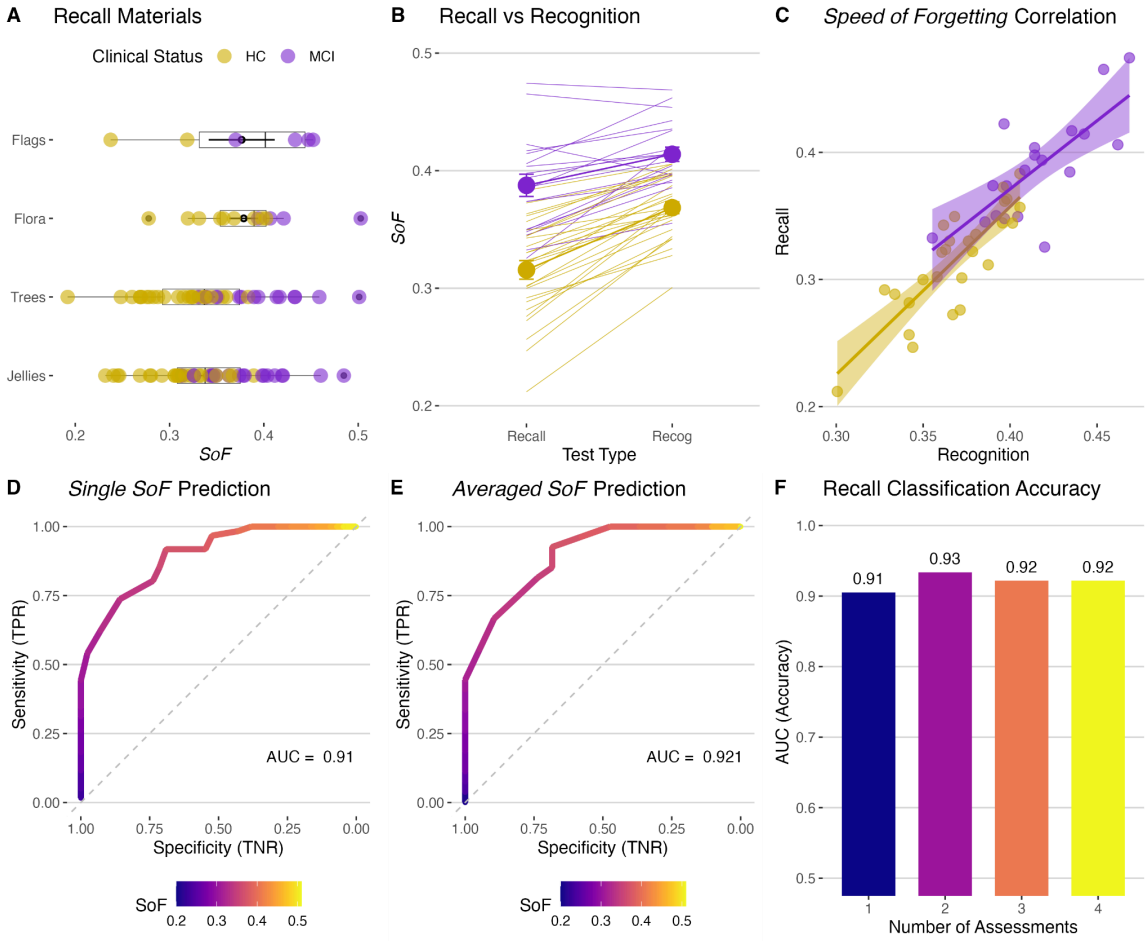

Figure 4. Comparative Analysis of Recall and Recognition Memory Performance.
(A) Recall Materials. Boxplots contrasting individual performance on memory tasks, stratified by lesson topic and clinical status (MCI in purple, HC in yellow), horizontal axis represents *Speed of Forgetting* (SoF). (B) Recall vs Recognition. Lines connect participant mean recall *SoF* versus mean recognition *SoF*. (C) S*peed of Forgetting* Correlation. Scatter plot depicting the relationship between recognition and recall performance, stratified by clinical status, shaded area represents 95% confidence interval. (D, E) ROC curves assessing the single- and average-measurement diagnostic accuracy of *SoF* for cognitive status, with the area under the curve (AUC) reflecting model precision. (F) Bar graphs depicting AUC values for multiple ROC curves, demonstrating the predictive accuracy based on the number of *SoF* assessments.