Routledge
Taylor & Francis Group

Journal of Clinical and Experimental Neuropsychology

ISSN: 1380-3395 (Print) 1744-411X (Online) Journal homepage: www.tandfonline.com/journals/ncen20

Recognition memory as a neuropsychological marker of progression in early-stage Alzheimer's disease: a 24-month prospective cohort study

María Julieta Russo, José Bueri, Lucía Alba-Ferrara, Gustavo Sevlever & Ricardo F. Allegri

To cite this article: María Julieta Russo, José Bueri, Lucía Alba-Ferrara, Gustavo Sevlever & Ricardo F. Allegri (16 Sep 2026): Recognition memory as a neuropsychological marker of progression in early-stage Alzheimer's disease: a 24-month prospective cohort study, Journal of Clinical and Experimental Neuropsychology, DOI: 10.1080/13803395.2026.2732062

To link to this article: https://doi.org/10.1080/13803395.2026.2732062

Published online: 16 Sep 2026.

Submit your article to this journal

Article views: 26

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ncen20

JOURNAL OF CLINICAL AND EXPERIMENTAL NEUROPSYCHOLOGY
https://doi.org/10.1080/13803395.2026.2732062

Routledge
Taylor & Francis Group

Check for updates

Recognition memory as a neuropsychological marker of progression in early-stage Alzheimer's disease: a 24-month prospective cohort study

María Julieta Russo[a,b], José Bueri[b], Lucía Alba-Ferrara[b], Gustavo Sevlever[a] and Ricardo F. Allegri[a,c]

[a]Centro de Memoria y Envejecimiento, INEU Fleni-CONICET, Buenos Aires, Argentina; [b]Hospital Universitario Austral, Facultad de Ciencias Biomédicas, Universidad Austral, Buenos Aires, Argentina; [c]Centro de Memoria y Envejecimiento, Fleni, Buenos Aires, Argentina

ABSTRACT

Introduction: Alzheimer's disease (AD) is the leading cause of dementia, and amnestic mild cognitive impairment (a-MCI) is considered its prodromal stage. Recognition memory is commonly assessed with yes/no tasks, although conventional scoring often overlook false alarms. Signal Detection Theory (SDT) measures, particularly discriminability (d-prime), may better detect early recognition deficits and predict progression to AD dementia.

Method: In this prospective cohort study, 278 older adults with a-MCI were followed for 24 months. Participants underwent clinical and neuropsychological assessment, including the Rey Auditory Verbal Learning Test. Recognition was decomposed into hits, false alarms, and d-prime. A subsample underwent amyloid and tau biomarker assessment by cerebrospinal fluid and/or imaging. Mixed-effects models examined longitudinal change and prediction of conversion to AD dementia.

Results: Delayed recall, recognition hits, and d-prime declined significantly over time, with steeper trajectories among converters. Converters showed an early and marked d'prime reduction during the first 12 months, mainly related to increased false alarms. Lower baseline d-prime and greater early decline independently predicted conversion after adjustment for demographic, global cognition, and AD biomarkers.

Conclusions: SDT-derived recognition indices, especially d'prime, provide clinically relevant prognostic information in a-MCI. Adding d' to delayed recall and biomarkers may improve risk stratification, follow-up planning, and selection for disease-modifying trials.

ARTICLE HISTORY
Received 3 March 2026
Accepted 30 August 2026

KEYWORDS
Episodic memory; recognition memory; signal detection theory; mild cognitive impairment; Alzheimer's disease; biomarkers

Introduction

Alzheimer's disease (AD) is the most common cause of dementia in older adults and represents a growing international public health challenge (Alzheimer's Association, 2023). Recent advances in biomarker identification and neuropsychological assessment have improved early diagnosis, particularly during the amnestic mild cognitive impairment (a-MCI) stage, considered the clinical precursor to AD dementia (Albert et al., 2011; Dubois et al., 2014).

Amnestic MCI primarily manifests as deficits in recent episodic memory, with patients experiencing notable difficulties in encoding and retrieving new information, both in free recall and recognition tasks (Petersen, 2004; Petersen et al., 2014). These memory impairments are associated with medial temporal lobe (MTL) dysfunction, particularly hippocampal pathology, which represents the earliest structural change in AD (Braak & Braak, 1991; C. R. Jack et al., 2010). Although this classic amnestic profile is characteristic of AD, substantial clinical and evolutionary heterogeneity exists among patients with a-MCI, making it difficult to predict which individuals will progress to dementia (Mitchell & Shiri-Feshki, 2009; Russo et al., 2017).

Annual conversion rates from a-MCI to AD dementia range from 10% to 15% (Petersen et al., 2001), but considerable individual variability exists. Not all patients with a-MCI harbor underlying AD pathology, and not all will progress to dementia within relatively short follow-up periods (Albert et al., 2011); some have non-AD or mixed etiologies and may remain clinically stable for several years. This heterogeneity underscores the critical need for improved prognostic markers that can identify high-risk subgroups early in the disease trajectory.

Episodic memory is traditionally assessed using instruments such as the Rey Auditory Verbal Learning Test (RAVLT), which measures learning, delayed recall, and recognition (Schmidt, 1996; Strauss et al., 2006). However, the conventional "one test-one domain"

CONTACT María Julieta Russo jrusso@fleni.org.ar; mariajulietarusso@gmail.com Centro de Memoria y Envejecimiento, INEU Fleni-CONICET, Montañeses 2325, C1428AQK, Buenos Aires, Argentina

© 2026 Informa UK Limited, trading as Taylor & Francis Group

approach has significant limitations due to the complexity of memory processes and the influence of attentional, executive, and emotional factors (Bondi et al., 2014; Edmonds et al., 2015).

Furthermore, in yes-no recognition tasks, failure to consider false alarms (incorrect endorsement of distractors) may underestimate the true degree of memory impairment (Snodgrass & Corwin, 1988; Wixted, 2007). Many neuropsychological tests provide normative data only for correct responses (hits), ignoring false alarm rates. This scoring system can be misleading because patients with severe memory disorders often exhibit a strong tendency to respond "yes," resulting in high hit rates but also elevated false alarms -a pattern that conventional scoring fails to capture adequately (Russo et al., 2013; Russo, Cohen, Campos, et al., 2016).

Signal Detection Theory (SDT) provides a mathematically rigorous framework for analyzing recognition memory that addresses these limitations (Green & Swets, 1966; Macmillan & Creelman, 2004). SDT-derived indices, such as discriminability (d-prime or d') and response bias (criterion, c), correct for individual response tendencies and capture the true ability to distinguish between studied items and distractors (Snodgrass & Corwin, 1988; Stanislaw & Todorov, 1999).

The discriminability index d-prime represents the distance between the signal (studied items) and noise (distractors) distributions in standardized units, reflecting memory sensitivity independent of response bias, whereas response bias (criterion, c) indexes the general tendency to endorse items as "old" or "new," reflecting pure memory sensitivity independent of response bias (Macmillan & Creelman, 2004). Together, these dimensionless SDT statistics enable objective comparisons of recognition performance across individuals and longitudinal time points (Russo, Cohen, Campos, et al., 2016).

Recognition memory is supported by two dissociable processes: familiarity (a rapid, automatic sense of prior encounter) and recollection (slower, deliberate retrieval of contextual details) (Mandler, 1980; Yonelinas, 2002). Dual-process models suggest that familiarity relies on perirhinal and entorhinal cortex, while recollection depends on hippocampal function (Aggleton & Brown, 1999; Eichenbaum et al., 2007). In early AD, hippocampal pathology impairs both recollection and, subsequently, familiarity-based recognition, resulting in deficits in both free recall and recognition (Russo et al., 2026; Wolk et al., 2008, 2011).

The d-prime index, while based on a single-process SDT model, effectively captures the overall discriminability that reflects the combined contribution of familiarity and recollection processes (Yonelinas, 2002). This pragmatic approach is clinically feasible and does not require complex experimental paradigms needed to separate dual processes (Russo, Cohen, Campos, et al., 2016). Previous studies have demonstrated that d-prime is sensitive to MTL dysfunction and predicts progression from a-MCI to AD dementia (Russo et al., 2017; Wolk et al., 2011).

The integration of AD biomarkers – amyloid-beta (Aβ) deposition, tau pathology, and neurodegeneration (the ATN framework) – has revolutionized AD diagnosis and research (Dubois et al., 2021; C. Jack et al., 2018). Cerebrospinal fluid (CSF) biomarkers, including decreased Aβ42 and elevated tau proteins (t-tau and p-tau181), are considered in vivo indicators of AD-related pathophysiology and, at the group level, are associated with an increased likelihood of progression from a-MCI to dementia (Blennow & Zetterberg, 2018; Hansson et al., 2018). There is now converging evidence that AD–related pathophysiological changes, as captured by CSF Aβ and tau concentrations and structural and functional neuroimaging, emerge many years before the onset of overt cognitive symptoms or a formal diagnosis of MCI (Greenberg et al., 2022; Jia et al., 2024; Moghekar et al., 2013). This temporal dissociation between biomarker abnormalities and clinical manifestations, formalized in the ATN research framework (C. R. Jack et al., 2016, 2024), underscores that biological markers alone may not fully capture the functional impact of underlying pathology at the prodromal stage. In this context, sensitive neuropsychological measures, particularly episodic recognition memory indices such as d-prime, are uniquely positioned to index the behavioral expression of early AD-related changes in a-MCI and to complement biomarker information for risk stratification and prognosis.

In addition, biomarker assessment is not universally accessible due to cost, invasiveness, and limited availability in low- and middle-income settings (Allegri et al., 2020; Russo, Cohen, Méndez, et al., 2016). This constraint further underscores the need for sensitive, low-cost neuropsychological markers that can be implemented at scale in routine practice (Bondi et al., 2014). In this pragmatic context, validating SDT-derived recognition indices such as d-prime becomes especially relevant for risk stratification and prognostic decision-making in memory clinics.

Despite decades of research on episodic memory and AD, few studies have systematically examined the prognostic value of SDT-derived recognition memory indices in a-MCI, particularly in longitudinal cohorts with biomarker data. The influence of false alarms on recognition memory profiles and their contribution to

conversion risk prediction has received limited attention (Russo et al., 2013; Russo, Cohen, Campos, et al., 2016).

This study addresses these gaps by examining recognition memory performance, decomposed into RAVLT delayed recall, recognition hits, and the SDT-derived discriminability index d-prime, as neuropsychological markers of progression to AD dementia in a 24-month prospective cohort of patients with a-MCI. Specifically, our primary aim was to determine whether baseline d-prime and its early longitudinal change independently predict conversion to AD dementia, beyond traditional recall measures and CSF AD biomarker status. As secondary aims, we sought to (a) compare the prognostic value of d-prime with other RAVLT indices and global cognitive and functional measures, (b) characterize the trajectories of RAVLT delayed recall, recognition hits, and d-prime over 12 and 24 months using linear mixed-effects models, and (c) explore how baseline recognition impairment and AD biomarker profiles jointly relate to cognitive decline risk. In this study, we focused on d-prime as the primary SDT-derived index of recognition discriminability, whereas response bias (c) was examined only in exploratory analyses.

Methods

Study design

The Study of Recognition Memory as a Neuropsychological Marker of Progression in Early-Stage Alzheimer's Disease (EMeR-Alz) was an exploratory, longitudinal, observational, prospective, non-randomized study conducted over 24 months. The study was approved by the Institutional Ethics Committee of Universidad Austral (approval number 005/2010), and all participants provided written informed consent prior to enrollment.

Participants

Between July and December 2020, 330 consecutive Spanish-speaking adults aged ≥ 60 years with a clinical diagnosis of a-MCI were recruited from the Memory and Aging Center at Instituto Fleni, Buenos Aires, Argentina. This cohort reflects older Spanish-speaking adults seen in a tertiary memory clinic in a middle-income Latin American country, where access to CSF and neuroimaging biomarkers is limited in routine practice. Participants underwent comprehensive baseline (T0), 12-month (T1), and 24-month (T2) evaluations. A total of 278 participants completed all three evaluations and constituted the final study sample.

Inclusion criteria

- Age ≥ 60 years.
- Clinical diagnosis of a-MCI according to established criteria (Albert et al., 2011; Petersen, 2004): (a) subjective memory complaint corroborated by an informant, (b) objective memory impairment on neuropsychological testing, (c) Mini-Mental State Examination (MMSE) score 24–30, (d) Clinical Dementia Rating (CDR) = 0.5, (e) preserved basic activities of daily living, (f) no dementia.
- Completion of baseline and 24-month follow-up assessments.

Exclusion criteria

- Any other neurological disorder besides the condition under study.
- Major psychiatric disorders.
- Severe or decompensated medical illness.
- Substance abuse or excessive alcohol consumption.
- Diagnosis of dementia at baseline.

Sample size and statistical power

Based on published literature indicating annual conversion rates of approximately 15% from MCI to dementia (Petersen et al., 2001), we estimated a 30% conversion rate over 24 months. Sample size calculations indicated that 250 participants would provide 95% confidence with a maximum acceptable error of 3%, adjusted for 15% attrition. To account for potential losses during follow-up, we initially enrolled 330 participants meeting all inclusion criteria.

Procedures

All participants underwent standardized neurological and neuropsychological evaluations at baseline (T0) and follow-up assessments at 12 months (T1) and 24 months (T2). At each visit, clinical characteristics were prospectively recorded. Consensus diagnosis was established at each timepoint by a multidisciplinary team including neurologists and neuropsychologists, based on clinical history, neurological examination, and neuropsychological test performance. Diagnostic criteria for a-MCI followed Petersen (2004) and Albert et al. (2011), and criteria for AD dementia followed National Institute on Aging–Alzheimer's Association (NIA-AA) guidelines (McKhann et al., 2011).

At the 24-month evaluation (T2), clinical diagnosis was reassessed and participants were classified as

converters (progressed to AD dementia) or non-converters (remained with MCI or reverted to normal cognition). Participants did not receive monetary compensation for study participation.

Cognitive and functional assessment

A comprehensive neuropsychological battery adapted from the Uniform Data Set (UDS) for Spanish-speaking populations was administered to characterize global episodic memory and other domains (Weintraub et al., 2009). In line with the aims of this study, the RAVLT provided the primary episodic memory indices under investigation. Additional measures (logical memory, language, attention/working memory, executive function, and mood) were included to describe the broader cognitive profile and to control for potential confounders, but they were not the focus of the prognostic analyses.

All neuropsychological test scores were converted to age- and education-adjusted z-scores using external Spanish-language normative data (Acevedo et al., 2009). Z-scores were computed such that the normative sample had a mean of 0 and a standard deviation of 1.5. Accordingly, positive z-scores indicate performance above the normative mean, values around 0 reflect average performance, and negative z-scores indicate performance below the normative mean, with more negative values indicating greater impairment.

Global cognition and functional status

- Mini-Mental State Examination (MMSE) (Folstein et al., 1975).
- Clinical Dementia Rating (CDR) (Morris, 1993).
- Functional Assessment Questionnaire (FAQ) (Pfeffer et al., 1982).

Memory

- Logical Memory (Story A) from Wechsler Memory Scale-Revised (WMS-R): immediate and delayed recall (Wechsler, 1987).
- RAVLT: total learning (sum of trials 1–5), RAVLT delayed recall (trial 7), recognition hits (trial 8), false alarms (Rey, 1964; Schmidt, 1996).

Language

- Semantic fluency (animals in 60 seconds).
- Boston Naming Test (30-item version).

Attention and working memory

- Digit Span Forward and Backward (Wechsler Adult Intelligence Scale-III).

Executive function and processing speed

- Trail Making Test Parts A and B (TMT-A, TMT-B) (Reitan, 1958).

Mood

- Geriatric Depression Scale (GDS) (Yesavage et al., 1983).

Episodic memory assessment and signal detection theory indices

The RAVLT was the primary instrument for assessing episodic memory (Rey, 1964; Schmidt, 1996). Standard measures included:

- Total learning: Sum of correctly recalled words across trials 1–5 (range: 0–75).
- Delayed recall: Number of words recalled after a 20–30-minute delay (range: 0–15).
- Recognition hits: Number of target words correctly identified in a yes-no recognition task (range: 0–15).
- False alarms: Number of distractors incorrectly identified as targets. For RAVLT false alarms, scores could range from 0 to 30, with higher values indicating more incorrect recognitions. This allowed us to characterize not only recognition hits but also error-prone response tendencies.

Signal detection theory indices

Following established SDT methodology (Macmillan & Creelman, 2004; Snodgrass & Corwin, 1988; Stanislaw & Todorov, 1999), we calculated:

- Discriminability (d-prime, d'): Discriminability (d-prime, d′) reflects the ability to distinguish studied items from distractors and is calculated as Z (Hit Rate) – Z (False Alarm Rate). Lower d-prime values indicate impaired discriminability, meaning that the participant has difficulty distinguishing previously studied words from non-studied distractors. In practice, this pattern is characterized by a combination of fewer correct recognitions (hits) and an increased rate of false alarms. In this study, d-prime values in the a-MCI sample typically ranged from approximately −2 to +1, with more negative scores indicating more severe impairment in recognition discriminability.
- Response bias (criterion, c): Indicates the participant's tendency to respond "yes" or "no," calculated as:

$$c = -0.5 \times [Z(\text{Hit Rate}) + Z(\text{False Alarm Rate})]$$

Values of c > 0 indicate a conservative bias (tendency to respond "no"), while c < 0 indicates a liberal bias (tendency to respond "yes").

In line with the study aims, d-prime was selected as the main SDT-derived index for longitudinal and prognostic analyses, because it simultaneously incorporates information from hits and false alarms and provides a direct measure of recognition discriminability. Response bias (c) was calculated for completeness and examined in exploratory analyses but was not included as a primary predictor in the main models.

Cerebrospinal fluid biomarkers

CSF samples were obtained by lumbar puncture at baseline, within ±3 months of the T0 clinical and neuropsychological assessment, in a subset of participants (approximately 50% of the sample) who consented to the procedure. CSF was collected in polypropylene tubes, centrifuged, aliquoted, and stored at −80°C until analysis. Biomarker concentrations were measured using commercially available enzyme-linked immunosorbent assay (ELISA) kits following manufacturer protocols:

- Amyloid-beta 1–42 (Aβ42)
- Total tau (t-tau)
- Phosphorylated tau at threonine 181 (p-tau181)

Biomarker positivity was defined according to established cutoff values (Blennow & Zetterberg, 2018): Aβ42 < 240 pg/mL, t-tau > specific threshold based on normative data, and p-tau181 > specific threshold. An AD biomarker profile was defined as Aβ42 < 240 pg/mL and ratio (t-tau / Aβ42) > 1.18 or (p-tau181 / Aβ42) > 1.3.

Statistical analysis

Statistical analyses were performed using SPSS version 23 (IBM Corporation, Armonk, NY, U.S.A.). Continuous variables were expressed as mean ± standard deviation (SD) and categorical variables as frequencies and percentages. Normality of continuous variables was assessed using the Kolmogorov–Smirnov test and inspection of histograms and Q–Q plots. As expected in a large clinical cohort, some variables showed statistically significant deviations from normality; however, distributions were approximately symmetric and without extreme floor or ceiling effects. Sensitivity analyses using non-parametric tests for group comparisons yielded the same pattern of significant and non-significant results as the parametric analyses. Given the robustness of parametric tests to mild deviations from normality in large samples and their greater statistical power, we retained parametric tests and linear mixed-effects models for the primary analyses. Statistical significance was set at $p < .05$ (two-tailed).

Baseline group comparisons

Baseline differences between converters and non-converters in demographic variables, global cognition, and neuropsychological measures were examined using *t*-tests for continuous variables and chi-square tests for categorical variables. Effect sizes for group differences were estimated using *Cohen's d* (standardized mean difference) computed from pooled standard deviations.

Longitudinal evolution of verbal memory and recognition

Repeated measures analysis of variance (ANOVA) was conducted to examine changes in cognitive performance across the three timepoints (T0, T1, T2) in the full sample. Time × Group (converter vs. non-converter) interactions were evaluated to assess differential trajectories. To account for individual variability and missing data, linear mixed-effects models were fitted with time (continuous), group (converter vs. non-converter), and their interaction as fixed effects, and participant as a random effect. Models controlled for baseline age, sex, education, and biomarker status (when available). Separate models were constructed for RAVLT delayed recall, recognition hits, false alarms, and d-prime as dependent variables. In sensitivity analyses, age, MMSE, and FAQ scores at baseline were entered as covariates to examine whether they altered the associations between RAVLT measures, amyloid status, and conversion.

To complement the repeated-measures ANOVAs, we fitted linear mixed-effects models using restricted maximum likelihood to examine predictors of longitudinal change in RAVLT delayed recall. Models specified subject-level random intercepts and a first-order factor-analytic (FA1) covariance structure for the repeated factor (time), and included conversion status, recognition hits, and d-prime as fixed effects in the primary model. In a second model, amyloid status was added as an additional fixed effect. To account for between-subject variability in both baseline performance and the strength of the association between delayed recall and d-prime, we also included subject-specific random components for d-prime. Amyloid status (CSF amyloid-positive vs. amyloid-negative) was included as a predictor in

supplementary mixed-effects models to examine whether recognition discriminability contributed to longitudinal memory trajectories beyond biomarker-defined Alzheimer pathology. CSF amyloid positivity is a core biological marker of AD, and its inclusion allowed us to test the incremental prognostic value of d-prime over and above amyloid burden.

For categorical predictors, non-converters served as the reference group for conversion status, and amyloid-negative status was considered the reference category in the biomarker subsample.

Recognition-based risk stratification

Participants were stratified by baseline RAVLT recognition hits performance into normal (z-score ≥ -1.5 SD) and impaired (z-score < -1.5 SD). This threshold was chosen because a deficit of at least 1.5 SD below age- and education-adjusted norms is widely used to define clinically meaningful impairment in episodic memory and to operationalize MCI in actuarial, multi-test neuropsychological criteria (Bondi et al., 2014; Edmonds et al., 2015). Group differences in conversion rates, cognitive trajectories, and biomarker profiles between these recognition-defined subgroups were examined using chi-square tests for categorical variables, t-tests or ANOVA for continuous variables, and the same mixed-effects procedures described above for longitudinal trajectories.

Results

Baseline characteristics

The final sample comprised 278 participants with a-MCI (mean age: 72.3 ± 6.8 years; 58% female; mean education: 12.4 ± 4.2 years). Baseline MMSE score was 26.7 ± 2.1, and all participants had CDR = 0.5. CSF biomarker data were available for 147/278 participants (52.9% of sample) and biomarker positivity (75.5%) was more frequent among converters than non-converters. The CSF subgroup was comparable to the full cohort in terms of age, sex distribution, years of education, and baseline MMSE and FAQ scores, indicating that participants who underwent lumbar puncture were broadly representative of the overall sample (data not shown).

At 24-month follow-up (T2), 104 participants (37.4%) had converted to AD dementia (converters), whereas 174 (62.6%) did not convert and either remained with a-MCI or reverted to normal cognition (non-converters). Among non-converters, 163 participants (58.6% of the total sample) remained with a-MCI and 11 (4.0% of the total sample) reverted to normal cognition at 24 months. Exploratory comparisons indicated that these two subgroups had broadly similar baseline demographic and neuropsychological profiles, so they were retained as a single non-converter category for the main analyses. Baseline demographic and clinical characteristics are presented in Table 1.

Baseline cognitive performance

At baseline, converters and non-converters differed significantly in global cognition and episodic memory performance. Converters showed lower MMSE scores, greater functional impairment on the FAQ, and poorer performance on RAVLT indices than non-converters (Table 2). Specifically, converters had reduced total learning across acquisition trials, lower delayed recall, and fewer recognition hits compared with non-converters.

With respect to error-prone recognition, converters also showed numerically higher RAVLT false alarm rates than non-converters; however, this difference did not reach statistical significance (Table 2). Accordingly, the marked reduction in d-prime values among converters was driven by the combined pattern of lower hit rates and a general tendency toward more incorrect endorsements, rather than by false alarms alone. As expected from the definition of d-prime, these differences indicate that converters had substantially impaired recognition discriminability—i.e., reduced

Table 1. Baseline demographic and clinical characteristics by conversion status at 24 months.

Variable	Non-converters ($n = 174$)	Converters ($n = 104$)	t / χ^2	p-value
Age (years)	75.28 ± 6.97	75.29 ± 6.51	−0.008	.994
Male sex, n (%)	114 (65.52)	66 (63.46)	0.121	.413
Education (years)	15.85 ± 2.91	15.58 ± 3.01	0.764	.446
MMSE	27.42 ± 1.71	26.65 ± 1.66	3.650	<.001
FAQ	2.49 ± 3.66	5.54 ± 4.86	−5.902	<.001
Biomarker-positive, n (%)	62/94 (65.96)	49/53 (92.45)	12.866	<.001

Abbreviatures: n, number of participants; X, mean; SD, standard deviation; MMSE, Mini Mental State Examination; FAQ, Functional Assessment Questionnaire.

Continuous variables are presented as X ± SD, and categorical variables as frequency (n) and percentage (%). AD biomarker data were available for 147 participants (94 non-converters and 53 converters). The p-value column indicates the result of the statistical comparison between groups, performed using Student's t-test or the chi-square test, depending on the nature of the data analyzed.

Table 2. Baseline neuropsychological performance by conversion status at 24 months.

Variable	Non-converters ($n = 174$)	Converters ($n = 104$)	t / χ^2	p-value	Cohen's d
ML – Immediate	7.89 (2.92)	6.41 (3.08)	3.997	.000	0.50
ML – Delayed	4.59 (2.69)	2.74 (2.21)	5.914	.000	0.73
RAVLT Total Learning	34.12 (9.81)	26.77 (6.17)	6.868	.000	0.85
RAVLT Delayed Recall	3.79 (3.77)	1.42 (1.86)	5.984	.000	0.74
RAVLT Recognition Hits	10.68 (3.34)	8.63 (3.87)	4.666	.000	0.58
RAVLT FA	1.74 (2.06)	2.07 (2.63)	−1.149	.252	−0.14
d-prime	2.26 (0.94)	1.76 (0.80)	4.534	.000	0.56
c	0.47 (0.38)	0.67 (0.51)	−3.648	.000	−0.46
Direct Span Digits	8.27 (2.03)	8.25 (2.03)	0.080	.936	0.01
Indirect Span Digits	6.63 (2.13)	5.99 (1.82)	2.565	.011	0.32
Boston Naming Test	26.10 (3.89)	24.95 (4.33)	2.289	.023	0.28
Semantic Fluency (Animals)	16.66 (4.98)	15.12 (4.54)	2.587	.010	0.32
TMT-A (seconds)	39.60 (14.94)	49.59 (26.74)	−3.996	.000	−0.50
TMT-B (seconds)	110.58 (59.75)	149.66 (75.89)	−4.748	.000	−0.59

Abbreviatures: n, number of participants; X, mean; SD, standard deviation; ML, Logical Memory; RAVLT, Rey Auditory Verbal Learning Test; FA, False Alarms; TMT, Trail Making Test.

Continuous variables are presented as X ± SD, and categorical variables as frequency (n) and percentage (%). The p-value column indicates the result of the statistical comparison between groups, performed using Student's t-test or the chi-square test, depending on the nature of the data analyzed. Effect sizes are Cohen's d (non-converters minus converters) based on pooled standard deviations. Positive d values indicate better performance in non-converters. For TMT-A and TMT-B (completion time in seconds), higher scores indicate slower performance; therefore, negative d values also reflect better performance in the non-converter group. For SDT indices, d' reflects discriminability (higher values = better recognition performance), whereas c reflects response bias (higher values = more conservative tendency to respond "old"). Positive d values indicate higher d' (better discriminability) in non-converters, whereas negative values for c indicate a more conservative response criterion in converters.

ability to distinguish previously studied words from distractors – already at the a-MCI stage.

Exploratory analyses of response bias (criterion c) showed that converters tended to exhibit a more liberal bias (greater tendency to answer "yes"), whereas non-converters showed a more conservative or neutral response style. However, group differences in c were largely explained by the underlying pattern of hits and false alarms, and response bias did not provide additional information beyond d-prime in accounting for baseline group differences. In line with the study aims, d-prime was therefore retained as the primary recognition index in subsequent longitudinal and stratification analyses, whereas c was examined only in exploratory models.

Longitudinal evolution of verbal memory

This subsection presents changes in verbal memory and recognition measures across the entire sample, with additional contrasts between converters and non-converters where indicated. We first examined mean trajectories of RAVLT delayed recall, recognition hits, and d-prime at baseline (T0), 12 months (T1), and 24 months (T2) in the full a-MCI cohort. Across all three indices, performance declined significantly over time (Table 3). Delayed recall worsened over 24 months, as did recognition hits, and d-prime showed the largest effect size, indicating that recognition discriminability declined more strongly than conventional recall and recognition hits in the overall sample.

To characterize the timing of decline, individual change scores were calculated for T0–T1 and T1–T2 for each RAVLT variable and paired-samples t-tests were conducted to compare the magnitude of change across intervals. For RAVLT delayed recall ($M = 0.05$, SD = 3.77; $t(277) = 0.21$, $p = .836$) and recognition hits ($M = 0.09$, SD = 5.86; $t(276) = 0.25$, $p = .806$), the magnitude of decline was similar across both time intervals, suggesting a relatively steady deterioration in these

Table 3. Repeated-measures ANOVA for verbal memory indices (total sample and recognition subgroups).

Variable	T0 X (SD)	T1 X (SD)	T2 X (SD)	F (df, error)	p-value	Partial η^2
Total sample ($n = 278$)						
RAVLT Delayed Recall	−1.81 (1.46)	−1.95 (1.55)	−2.08 (1.54)	10.24 (2,552)	<.001	0.036
RAVLT Recognition Hits	−1.17 (1.46)	−1.43 (1.68)	−1.66 (1.69)	16.56 (2,552)	<.001	0.057
d-prime	2.07 (0.93)	1.82 (0.92)	1.78 (1.02)	21.87 (2,552)	<.001	0.073
Normal baseline recognition hits ($n = 187$)						
RAVLT Delayed Recall	−1.17 (1.59)	−1.27 (1.72)	−1.41 (1.76)	10.24 (2,552)	<.001	0.036
RAVLT Recognition Hits	−0.27 (0.82)	−0.65 (1.39)	−0.82 (1.35)	16.56 (2,552)	<.001	0.057
d-prime	2.57 (0.82)	2.26 (0.92)	2.30 (1.02)	21.87 (2,552)	<.001	0.073
Impaired baseline recognition hits ($n = 91$)						
RAVLT Delayed Recall	−2.31 (0.92)	−2.21 (1.19)	−2.29 (1.00)	1.84 (2,176)	0.16	0.020
RAVLT Recognition Hits	−2.75 (0.68)	−2.23 (1.59)	−2.05 (1.29)	6.75 (2,176)	.001	0.071
d-prime	1.29 (0.59)	1.53 (0.86)	1.55 (0.70)	9.55 (2,176)	<.001	0.103

X, media; SD, standard deviation; df, degrees of freedom; RAVLT, Rey Auditory Verbal Learning Test.

Table 4. Linear mixed-effects models for without and with amyloid status results of linear mixed-effects models for RAVLT delayed recall without and with amyloid status as a fixed effect.

Effect	*B*	SE	*F*	*p-value*
Model 1: without amyloid				
Intercept	0.35	0.09	1371.94	<.001
Conversion at 24 months	0.19	0.01	25.11	<.001
RAVLT recognition (hits)	0.00	0.01	1.43	.232
d-prime	0.11	0.02	134.19	<.001
RAVLT recognition (hits)*d-prime	0.01	0.01	0.11	.744
Model 2: with amyloid				
Intercept	0.35	0.09	365.38	<.001
Conversion at 24 months	0.19	0.01	2.76	.100
RAVLT recognition (hits)	0.01	0.02	0.10	.752
d-prime	0.11	0.02	39.63	<.001
RAVLT recognition (hits)*d-prime	0.01	0.01	1.36	.246
Amyloid status (positive)	0.70	0.21	11.10	.001

Abbreviatures: B, unstandardized coefficient; SE, standard error of B; F, F statistic for the Type III fixed effect; p, probability value; RAVLT, Rey Auditory Verbal Learning Test.

The dependent variable is RAVLT delay recall (z score). Conversion status (CONVAD24), recognition, d-prime, and amyloid status are fixed-effect predictors. Model 1 includes conversion status, recognition, and d-prime as fixed effects. Model 2 additionally includes amyloid status. Both models specify subject-level random intercepts and random components for d-prime, with an FA1 covariance structure for time.

measures. In contrast, d-prime showed a significantly greater decrease between T0 and T1 than between T1 and T2 ($M = 0.21$, SD = 1.32; $t(276) = 2.60$, $p = .010$), indicating a more accelerated loss of discriminative sensitivity during the first year of follow-up.

We then examined these trajectories separately in converters and non-converters, as illustrated in Figure 1. In converters (Figure 1A), RAVLT delayed recall, recognition hits, and d-prime all declined significantly from baseline to 12 months and from 12 to 24 months (all $p < .01$), consistent with progressive deterioration in verbal episodic memory and recognition discriminability over time. In non-converters (Figure 1B), RAVLT delayed recall and recognition hits remained relatively stable (all $p > .10$), with no significant change across assessments, and d-prime showed only a small, isolated decrease between T0 and T1 ($p = .009$) that was not followed by further decline from T1 to T2 ($p = .525$). This relative stability in recognition discriminability among non-converters contrasts with the pronounced decline observed in converters and supports the role of d-prime as a longitudinal marker of progression rather than merely cross-sectional impairment.

A repeated-measures multivariate ANOVA was conducted to examine longitudinal changes in verbal memory performance across three time points (baseline, 12 months, and 24 months) and to test whether trajectories differed as a function of conversion to dementia at 24 months. The three primary indices were RAVLT delayed recall, RAVLT recognition hits, and d-prime. At the multivariate level, there was a significant main effect of time on overall verbal memory performance (Wilks' $\Lambda = 0.85$, $F(6, 271) = 7.97$, $p < .001$, partial $\eta^2 = .15$) and a significant main effect of conversion status (Wilks' $\Lambda = 0.81$, $F(3, 273) = 63.47$, $p < .001$, partial $\eta^2 = .19$), as well as a significant time * group interaction (Wilks' $\Lambda = 0.96$, $F(6, 270) = 3.50$, $p = .002$, partial $\eta^2 = .02$), indicating differential trajectories for converters and non-converters.

In the total a-MCI cohort, all three RAVLT scores showed significant longitudinal decline (Table 3). Delayed recall worsened over time ($F(2, 552) = 10.24$, $p < .001$, partial $\eta^2 = .036$), as did recognition hits scores ($F(2, 552) = 16.56$, $p < .001$, partial $\eta^2 = .057$). d-prime showed the largest effect size ($F(2, 552) = 21.87$, $p < .001$, partial $\eta^2 = .073$), indicating that recognition discriminability declined more strongly than conventional recall and recognition hits in the overall sample.

Recognition-based risk stratification

Participants were stratified according to baseline RAVLT recognition hits performance: 187 (67.27%) had normal RAVLT recognition hits (z-score ≥ -1.5 SD), and 91 (32.73%) had impaired RAVLT recognition hits (z-score < -1.5 SD). Conversion rates differed markedly between subgroups (45.0% vs. 22.2%; $\chi^2 = 15.8$, $p < .001$). Among participants with normal baseline RAVLT recognition hits, repeated-measures ANOVA showed significant longitudinal decline over time in delayed recall, recognition hits, and d-prime (all $p < .01$). In contrast, in participants with impaired baseline recognition hits, delayed recall remained relatively stable ($p = .16$), whereas recognition hits and d-prime changed significantly over time (both $p \leq .001$), again with d-prime showing the largest effect size. Across the

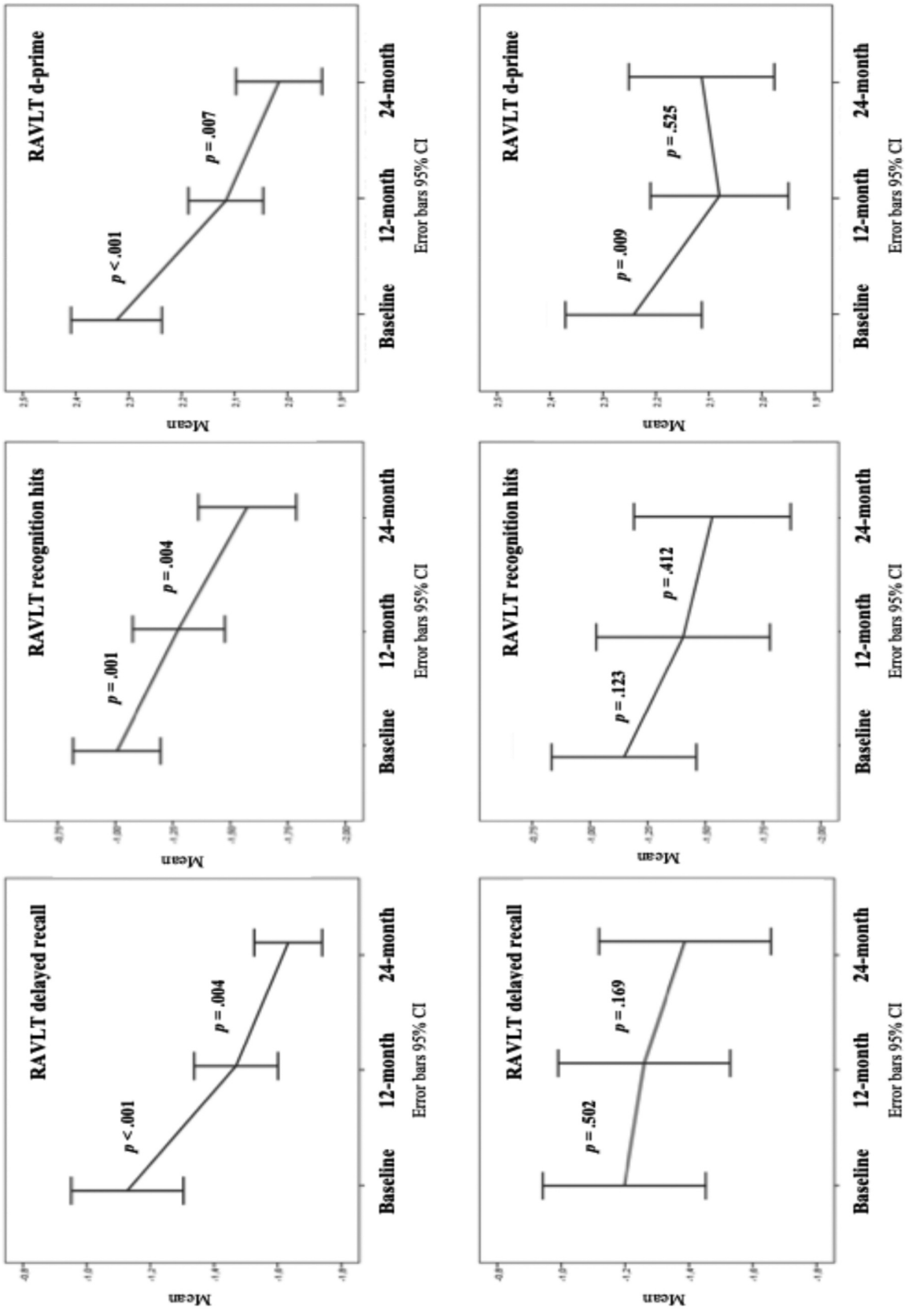

Figure 1. RAVLT performance in converters and non-converters. Longitudinal trajectories of RAVLT delayed recall, recognition hits, and d-prime in converters (A) and non-converters (B). The y-axes represent mean z-scores for each measure at baseline (T0), 12months (T1), and 24months (T2). Values below zero indicate performance below normative expectations, and more negative scores reflect greater impairment. Error bars represent 95% confidence intervals for the mean at each time point. Within each panel,*p*-values refer to paired comparisons between consecutive time points (T0–T1 and T1–T2). The plots highlight progressive decline in all three indices in converters, whereas non-converters show relative stability, with only a small, non-sustained decrease in d-prime between T0 and T1.

full sample and both recognition-defined subgroups, d-prime emerged as the most sensitive RAVLT index of longitudinal change.

In summary, a progressive and significant decline was observed in all three variables over the follow-up, which was more pronounced in the converter group. Among the memory measures, d-prime consistently showed the largest effect size and the most robust differences between groups and over time, both in the overall sample and within subgroups defined by normal or impaired baseline recognition memory (hits). These associations remained essentially unchanged after additional adjustment for baseline age, MMSE, and FAQ (data not shown).

Mixed linear models

In linear mixed-effects models (Table 4) with RAVLT delayed recall as the dependent variable, conversion status ($F(1,230) = 25.11$, $p < .001$) and d-prime ($F(1,552.7) = 134.19$, $p < .001$) were significant predictors of delayed recall performance over time, whereas recognition hits ($F(1,593.6) = 1.43$, $p = .232$) and the recognition hits * d-prime interaction ($F(1,292.7) = 0.11$, $p = .744$) were not. When amyloid positivity was added as an additional fixed effect, both d-prime ($F(1,234.48) = 39.63$, $p < .001$) and amyloid positivity ($F(1,143.53) = 11.10$, $p = .001$) remained significant, while conversion status, recognition hits, and the recognition hits * d-prime interaction remained non-significant (all $p > .05$). Overall, these models indicate that SDT-based discriminability (d-prime) and amyloid status are the main longitudinal predictors of delayed recall, whereas conventional recognition scores (hits) and conversion status do not explain additional variance once these factors are considered.

Discussion

This prospective, longitudinal study over a 24-month follow-up period demonstrates that recognition memory assessment, particularly the d-prime discriminability index derived from SDT, provides clinically valuable information for identifying patients with a-MCI at higher risk of progression to AD dementia. Our findings support four key conclusions: (1) RAVLT d-prime (recognition discriminability) declines early and steeply in patients who subsequently convert to dementia; (2) baseline RAVLT d-prime impairment and early decline predict conversion independently of biomarker status; (3) integration of neuropsychological and biomarker data optimizes prognostic accuracy; and (4) recognition memory assessment addresses limitations of conventional scoring methods that ignore false alarms.

Early and steep decline in d-prime as a prognostic marker

A central finding of this study is that the d-prime discriminability index showed pronounced decline during the first 12 months in converters, preceding the clinical diagnosis of dementia by an average of 12–18 months. This early decline suggests that loss of discriminability – the ability to distinguish studied items from novel distractors – is a sensitive and early marker of accelerated cognitive decline in a-MCI. The steep trajectory of RAVLT d-prime decline contrasts with the more gradual decline observed in delayed recall, suggesting that recognition memory processes may be particularly vulnerable during the transition from a-MCI to dementia. From a clinical perspective, a cumulative decline of approximately 1 SD in d-prime over 24 months is unlikely to reflect measurement noise alone and instead represents a meaningful worsening in recognition discriminability. Such a change is consistent with a shift from borderline to clearly impaired performance on standardized norms and may correspond to an increased frequency of everyday memory failures related to difficulty distinguishing previously encountered from novel information. In this context, the early and steep decline observed between T0 and T1 suggests that d-prime captures a critical window of accelerated deterioration that precedes formal dementia diagnosis and could be used to flag individuals who warrant closer monitoring and earlier intervention.

An important caveat is that diagnostic classification (converter vs. non-converter) was only reassessed at 24 months. As a result, some participants likely crossed the diagnostic threshold between T0 and T1 but were still categorized as converters at T2 for analytic purposes. This implies that the observed early decline in d-prime may underestimate the temporal proximity between loss of discriminability and the moment of clinical conversion, and that the trajectories reported here should be interpreted as approximations rather than precise time-locked markers of diagnosis.

This finding is consistent with neurobiological models proposing that recognition memory relies on both familiarity-based and recollection-based processes, mediated by partially dissociable neural substrates (Eichenbaum et al., 2007; Yonelinas, 2002). Familiarity, supported by perirhinal and entorhinal cortex, may be disrupted early in AD as neurofibrillary tangles spread from transentorhinal cortex (Braak & Braak, 1991). Recollection, dependent on hippocampal

integrity, is also severely compromised in AD (Eichenbaum et al., 2007; Yonelinas, 2002). The d-prime index, which reflects overall discriminability, captures the combined impairment of these processes and may therefore be behaviorally sensitive to Alzheimer-related pathological burden. In our sample, lower d-prime scores were associated with CSF amyloid positivity and with clinical conversion, supporting its potential utility as a neuropsychological proxy of early disease-related change rather than a direct biomarker. Consistent with our a priori focus, we did not model SDT response bias (criterion c) as a prognostic marker. While c captures decision tendencies that can vary partly independently from discriminability, our analyses were designed to evaluate d-prime as the main SDT-derived indicator of hippocampal-dependent recognition memory; future studies should examine whether longitudinal changes in response bias provide additional prognostic information beyond d-prime.

In the present a-MCI sample, these mechanisms are reflected in the clear separation between converters and non-converters: converters showed a steeper decline in d-prime and a tendency toward higher false alarm rates over time, suggesting a greater disruption of the ability to distinguish studied from novel items during the prodromal-to-dementia transition. Thus, the neurobiological models reviewed above are relevant not only to early AD in general, but specifically to the subgroup of a-MCI patients who subsequently progressed to dementia in this cohort.

Independent predictive value beyond biomarkers

A notable finding is that d-prime decline retained significant prognostic value even after adjusting for CSF biomarker status. Although biomarker-positive individuals had higher baseline risk, cognitive trajectories – particularly early d-prime decline – provided additional prognostic information, suggesting that neuropsychological measures capture dynamic disease progression and functional impact that static biomarkers may not fully reflect. This perspective aligns with emerging frameworks that emphasize the complementary roles of biological and cognitive markers in AD diagnosis and prognosis (Dubois et al., 2021; C. Jack et al., 2018) and underscores the importance of neuropsychological assessment in settings where biomarker access is limited or costly, as is often the case in Latin America and other resource-limited regions (Allegri et al., 2020).

In supplementary mixed-effects models with RAVLT delayed recall as the dependent variable, conversion status and d-prime – but not recognition hits or their interaction – significantly explained variance in delayed recall trajectories over time. When amyloid positivity was added, d-prime and amyloid positivity remained significant fixed effects, whereas recognition hits continued to be non-significant, indicating that d-prime captures prognostically relevant information beyond conventional recognition scores (hits) and that its contribution is independent and complementary to that of CSF amyloid biomarkers.

Taken together, these findings have several implications for the use of recall versus recognition-based indices as prognostic markers in a-MCI. Beyond its association with biomarker status, d-prime also showed greater sensitivity to longitudinal change than conventional memory indices. In the present sample, delayed recall and recognition hits declined over time but did not exhibit the same degree of prognostic specificity once other factors were considered, whereas d-prime remained closely aligned with conversion trajectories. The fact that conversion status no longer strongly predicted delayed recall in the updated analyses suggests that traditional recall measures may be more vulnerable to ceiling effects, practice effects, and compensatory strategies, which can obscure early disease-related changes. In contrast, recognition discriminability appears to provide a more direct behavioral index of subtle degradation in familiarity- and recollection-based processes, supporting its use as a complementary marker alongside delayed recall in prognostic assessment.

Addressing limitations of conventional recognition scoring

Our findings highlight a critical limitation of conventional recognition memory scoring: reliance solely on correct responses (hits) without accounting for false alarms. Converters in this study exhibited not only lower hit rates but also a tendency toward higher false alarm rates, resulting in reduced recognition discriminability. Although the difference in false alarm rates between converters and non-converters was not statistically significant at baseline, the overall pattern of more liberal endorsement and reduced d-prime in converters is consistent with impaired monitoring and source memory processes (Ally et al., 2013; Stark et al., 2015).

The d-prime index, by incorporating both hits and false alarms, corrects for individual response biases and provides a purer measure of memory sensitivity. This approach is particularly important in clinical populations where response tendencies may vary due to cognitive impairment, anxiety, or executive dysfunction (Russo, Cohen, Campos, et al., 2016; Snodgrass & Corwin, 1988). Our results validate the clinical utility of SDT-derived indices and support their routine

incorporation into neuropsychological assessment batteries for MCI.

Clinical implications and risk stratification

The present findings have important implications for clinical practice, clinical trial design, and personalized medicine in AD. The longitudinal nature of the results is clinically relevant: a marked decline in d-prime over the first 12 months may help identify patients with a-MCI who require closer follow-up, earlier biomarker work-up, and more proactive counseling regarding progression risk. In memory-clinic settings, such trajectory-based information is particularly valuable when formal diagnostic conversion has not yet occurred, but the pattern of cognitive change suggests accelerating disease progression.

Identification of high-risk subgroups based on recognition memory performance enables more targeted monitoring, earlier interventions, and optimized resource allocation. This is especially pertinent in Latin American and other resource-limited settings, where biomarker testing is not universally available and follow-up capacity is constrained; in these contexts, SDT-based recognition indices derived from standard tests such as the RAVLT offer a pragmatic, scalable tool to refine risk stratification without additional infrastructure. Patients with impaired baseline recognition hits (z-score < -1.5 SD) had nearly double the conversion rate compared to those with preserved recognition hits, suggesting that this threshold may be clinically useful for risk stratification.

Incorporating d-prime into routine practice is feasible, as it relies on standard recognition memory tasks (e.g., RAVLT, California Verbal Learning Test) and simple calculations. Clinicians can use freely available software or spreadsheets to compute d-prime from raw data, making this approach accessible even in resource-limited settings (Russo, Cohen, Campos, et al., 2016).

For clinical trials of disease-modifying therapies, enrichment strategies based on cognitive markers may improve trial efficiency by selecting participants at higher risk of progression. Our findings suggest that combining neuropsychological markers such as d-prime with biomarkers can optimize participant selection and potentially reduce sample size requirements and trial duration (Russo et al., 2017).

Neurobiological interpretation: Familiarity and discriminability

The pronounced impairment in d-prime observed in converters likely reflects disruption of familiarity-based recognition processes, which depend on perirhinal and entorhinal cortex – regions affected early in AD pathology (Braak & Braak, 1991; Juottonen et al., 1998). Reduced discriminability may also reflect impaired pattern separation, a hippocampal-dependent process that enables fine-grained distinctions between similar stimuli (Yassa & Stark, 2011). Animal models and human imaging studies have demonstrated that dentate gyrus and CA3 hippocampal subfields, critical for pattern separation, are particularly vulnerable in early AD (Bakker et al., 2012; Small et al., 2011).

High false alarm rates in converters may reflect a failure of memory monitoring and source memory, processes supported by prefrontal cortex and hippocampus (Simons & Spiers, 2003). As AD pathology spreads from MTL to association cortices, these executive and metacognitive processes become increasingly impaired, contributing to elevated false alarms and reduced d-prime (Gallo et al., 2006).

Heterogeneity in a-MCI and prognostic profiles

Our findings underscore the clinical and neuropsychological heterogeneity within the a-MCI diagnostic category. Not all participants with impaired RAVLT delayed recall showed impaired recognition, and not all converted to dementia. This variability reflects diverse underlying etiologies: some patients may have AD-related pathology, whereas others may have non-AD or mixed causes of memory impairment (e.g., vascular disease, depression, medication effects) (Albert et al., 2011; Petersen et al., 2014).

The subgroup with preserved recognition despite impaired delayed recall may represent individuals with retrieval deficits rather than true encoding/storage impairment – a profile more consistent with frontal-subcortical dysfunction or non-AD conditions (Bondi et al., 2014). Conversely, the subgroup with impaired recognition and high false alarms likely reflects greater involvement of medial temporal lobe structures, in a pattern that is consistent with underlying AD pathology and may help explain their higher conversion rates; however, given the presence of biomarker-negative converters, alternative or mixed etiologies cannot be excluded.

In the non-converter group, participants included both individuals with persistent amnestic MCI and those who reverted to normal cognition at follow-up. This heterogeneity may attenuate within-group differences in longitudinal trajectories of recognition memory, particularly for discriminability measures such as d-prime. Our decision to combine these subgroups was driven by the need to preserve statistical power and

maintain a parsimonious model, but this constraint should be considered when interpreting the stability of performance among non-converters.

Although the present results are centered on one memory measure, they support the broader principle that prognostic assessment in a-MCI benefits from combining complementary cognitive markers rather than relying on a single test score (Bondi et al., 2014; Edmonds et al., 2015). Comprehensive assessment capturing encoding, retrieval, recognition, and false alarm rates enables more accurate diagnostic classification and prognostic stratification.

Strengths and limitations

Strengths of this study include the prospective, longitudinal design with standardized assessments at three timepoints; the inclusion of CSF biomarkers in a substantial subset; the application of SDT methodology to derive d-prime indices; and the focus on a clinically relevant question with direct implications for practice and trial design. The sample size ($n = 278$) provides adequate statistical power for longitudinal analyses and multivariate modeling.

Limitations include the single-center design, which may limit generalizability to other populations, particularly non-Spanish-speaking or culturally diverse cohorts. The 24-month follow-up, while clinically meaningful, is relatively brief; longer follow-up would enable examination of more distant outcomes and cumulative conversion rates. Although we conducted sensitivity analyses adjusting for baseline age, MMSE, and functional status (FAQ), residual confounding by baseline disease severity cannot be fully excluded given the observational design.

In addition, diagnostic status was only reassessed at the 24-month visit, so the exact timing of conversion between T0 and T2 could not be determined, which may blur the temporal alignment between changes in d-prime and the formal diagnosis of dementia. A further limitation is that CSF biomarker data were available for only 50% of the sample, which reduces power for biomarker-stratified analyses and raises the possibility of selection bias; however, this subsample was broadly similar to the full cohort in terms of key demographic and clinical characteristics, partially mitigating this concern.

Another limitation is the potential etiological heterogeneity within the a-MCI sample. Only about 60% of participants who underwent CSF assessment were biomarker positive, and approximately 22% of converters were biomarker negative, indicating that non-AD or mixed pathologies likely contributed to the observed trajectories. As a result, the present findings should be interpreted as reflecting progression in an a-MCI cohort enriched for, but not restricted to, underlying AD pathology, and mechanistic inferences specific to AD should be made with caution.

Neuroimaging biomarkers (amyloid PET, tau PET, structural MRI) were not systematically available, precluding examination of structure-function relationships. Future studies should examine whether d-prime indices predict specific patterns of cortical atrophy or tau deposition, and whether interventions targeting memory processes can modify d-prime trajectories. Validation in independent cohorts, including diverse linguistic and cultural backgrounds, is also needed.

Future work should also examine response styles in more detail, including whether a positive ("yea-saying") response bias predicts conversion or changes with progression to dementia. Although response bias (criterion c) was only explored descriptively in the present study, a more systematic analysis of liberal versus conservative response tendencies could provide additional insight into how decision styles interact with recognition discriminability in a-MCI.

Conclusions

This study provides robust evidence that recognition memory assessment, particularly the d-prime discriminability index derived from SDT, is a valuable neuropsychological marker for identifying patients with a-MCI at high risk of progression to AD dementia. D-prime shows early and pronounced decline in converters, predicts conversion independently of biomarkers, and addresses critical limitations of conventional recognition scoring by accounting for false alarms. Integration of d-prime assessment with biomarker data optimizes prognostic accuracy and risk stratification.

These findings have direct clinical implications: routine incorporation of d-prime assessment into neuropsychological batteries is feasible, cost-effective, and improves diagnostic precision. By demonstrating the prognostic value of d-prime in a large Latin American a-MCI cohort, this study illustrates how advanced recognition memory indices can be integrated into routine assessment to support clinical decision-making in settings where access to costly biomarkers is limited. For clinical trials, enrichment strategies combining neuropsychological and biomarker criteria may enhance trial efficiency.

Future research should focus on validating these findings in diverse populations, examining neural correlates of d-prime decline, and exploring whether

cognitive interventions can modify recognition memory trajectories and delay progression to dementia.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Data availability statement

De-identified data that support the findings of this study are available from the corresponding author upon reasonable request, subject to approval by the local ethics committee and in accordance with institutional and national regulations on the protection of human participants' data.

References

Acevedo, A., Krueger, K. R., Navarro, E., Ortiz, F., Manly, J. J., Padilla-Vélez, M. M., Mungas, D., López, O. L., & Mungas, D. (2009). The Spanish translation and adaptation of the uniform data Set of the National institute on Aging Alzheimer's disease centers. *Alzheimer Disease & Associated Disorders*, *23*(2), 102–109. https://doi.org/10.1097/WAD.0b013e318193e376

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences*, *22*(3), 425–444. https://doi.org/10.1017/S0140525X99002034

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National institute on Aging-Alzheimer's Association workgroups. *Alzheimer's & Dementia*, *7*(3), 270–279.

Allegri, R. F., Chrem Méndez, P., Calandri, I., Cohen, G., Martín, M. E., Russo, M. J., Sevlever, G., Pertierra, L., Tapajóz, F., Clarens, M. F., Campos, J., Nahas, F. E., Vázquez, S., Surace, E., & Sevlever, G. (2020). Prognostic value of ATN Alzheimer biomarkers: 60-month follow-up results from the Argentine Alzheimer's disease neuroimaging initiative. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *12*(1), e12026. https://doi.org/10.1002/dad2.12026

Ally, B. A., Hussey, E. P., Ko, P. C., & Molitor, R. J. (2013). Pattern separation and pattern completion in Alzheimer's disease: Evidence of rapid forgetting in amnestic mild cognitive impairment. *Hippocampus*, *23*(12), 1246–1258. https://doi.org/10.1002/hipo.22162

Association, A. (2023). Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, *19*(4), 1598–1695.

Bakker, A., Krauss, G. L., Albert, M. S., Speck, C. L., Jones, L. R., Stark, C. E., Yassa, M., Bassett, S., Shelton, A., & Gallagher, M. (2012). Reduction of hippocampal hyperactivity improves cognition in amnestic mild cognitive impairment. *Neuron*, *74*(3), 467–474. https://doi.org/10.1016/j.neuron.2012.03.023

Blennow, K., & Zetterberg, H. (2018). Biomarkers for Alzheimer's disease: Current status and prospects for the future. *Journal of Internal Medicine*, *284*(6), 643–663. https://doi.org/10.1111/joim.12816

Bondi, M. W., Edmonds, E. C., Jak, A. J., Clark, L. R., Delano-Wood, L., McDonald, C. R., Salmon, D. P., Libon, D. J., Au, R., Galasko, D., & Salmon, D. P. (2014). Neuropsychological criteria for mild cognitive impairment improve diagnostic precision, biomarker associations, and progression rates. *Journal of Alzheimer's Disease*, *42*(1), 275–289. https://doi.org/10.3233/JAD-140276

Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica*, *82*(4), 239–259. https://doi.org/10.1007/BF00308809

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K., Cummings, J. L., Gauthier, S., Selkoe, D., Bateman, R., Cappa, S., Crutch, S., Engelborghs, S., Frisoni, G. B., Fox, N. C., Galasko, D., Habert, M.-O., Jicha, G. A., Nordberg, A. & Scheltens, P. (2014). Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurology*, *13*(6), 614–629. https://doi.org/10.1016/S1474-4422(14)70090-0

Dubois, B., Villain, N., Frisoni, G. B., Rabinovici, G. D., Sabbagh, M., Cappa, S., Feldman, H. H., Bombois, S., Epelbaum, S., Teichmann, M., Habert, M.-O., Nordberg, A., Blennow, K., Galasko, D., Stern, Y., Rowe, C. C., Salloway, S., Schneider, L. S., Cummings, J. L., & Feldman, H. H. (2021). Clinical diagnosis of Alzheimer's disease: Recommendations of the International working group. *Lancet Neurology*, *20*(6), 484–496. https://doi.org/10.1016/S1474-4422(21)00066-1

Edmonds, E. C., Delano-Wood, L., Clark, L. R., Jak, A. J., Nation, D. A., McDonald, C. R., Bondi, M. W., Au, R., Galasko, D., Salmon, D. P., & Bondi, M. W. (2015). Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimer's & Dementia*, *11*(4), 415–424. https://doi.org/10.1016/j.jalz.2014.03.005

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*(1), 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Gallo, D. A., Sullivan, A. L., Daffner, K. R., Schacter, D. L., & Budson, A. E. (2006). Associative recognition in Alzheimer's disease: Evidence for impaired recall-to-reject. *Neuropsychology*, *20*(1), 78–88.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Greenberg, B. D., Pettigrew, C., Soldan, A., Wang, J., Wang, M. C., Darrow, J. A., Albert, M. S., & Moghekar, A. (2022). CSF Alzheimer disease biomarkers: Time-varying relationships with MCI symptom onset and associations with age, sex, and ApoE4. *Neurology*, *99*(15), e1640–e1650. https://doi.org/10.1212/WNL.0000000000200953

Hansson, O., Seibyl, J., Stomrud, E., Zetterberg, H., Trojanowski, J. Q., Bittner, T., Lifke, V., Corradini, V., Eichenlaub, U., Batrla, R., Buck, K., Zink, K., Rabe, C., Blennow, K., & Shaw, L. M. (2018). CSF biomarkers of Alzheimer's disease concord with amyloid-β PET and predict clinical progression: A study of fully automated immunoassays in BioFINDER and ADNI cohorts. *Alzheimer's & Dementia*, *14*(11), 1470–1481. https://doi.org/10.1016/j.jalz.2018.01.010

Jack, C., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Karlawish, J., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M. &Ryan, L. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, *14*(4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

Jack, C. R., Jr., Andrews, J. S., Beach, T. G., Buracchio, T., Dunn, B., Graf, A., Hansson, O., Ho, C., Jagust, W., McDade, E., Molinuevo, J. L., Okonkwo, O. C., Pani, L., Rafii, M. S., Scheltens, P., Siemers, E., Snyder, H. M., Sperling, R., Teunissen, C. E., & Carrillo, M. C. (2024). Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *20*(8), 5143–5169.

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Feldman, H. H., Frisoni, G. B., Hampel, H., Jagust, W. J., Johnson, K. A., Knopman, D. S., Petersen, R. C., Scheltens, P., Sperling, R. A., & Dubois, B. (2016). A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology*, *87*(5), 539–547. https://doi.org/10.1212/WNL.0000000000002923

Jack, C. R., Knopman, D. S., Jagust, W. J., Shaw, L. M., Aisen, P. S., Weiner, M. W., Petersen, R. C., & Trojanowski, J. Q. (2010). Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurology*, *9*(1), 119–128. https://doi.org/10.1016/S1474-4422(09)70299-6

Jia, J., Ning, Y., Chen, M., Wang, S., Yang, H., Li, F., Ding, J., Li, Y., Zhao, B., Lyu, J., Yang, S., Yan, X., Wang, Y., Qin, W., Wang, Q., Li, Y., Zhang, J., Liang, F., Liao, Z., & Wang, S. (2024). Biomarker changes during 20 years preceding Alzheimer's disease. *The New England Journal of Medicine*, *390*(8), 712–722. https://doi.org/10.1056/NEJMoa2310168

Juottonen, K., Laakso, M. P., Insausti, R., Lehtovirta, M., Pitkänen, A., Partanen, K., & Soininen, H. (1998). Volumes of the entorhinal and perirhinal cortices in Alzheimer's disease. *Neurobiology of Aging*, *19*(1), 15–22. https://doi.org/10.1016/S0197-4580(98)00007-4

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide* (2nd ed.) Psychology Press. https://doi.org/10.4324/9781410611147 taylorfrancis

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*(3), 252–271. https://doi.org/10.1037/0033-295X.87.3.252

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C., Kawas, C. H., Phelps, C. H., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National institute on Aging-Alzheimer's Association workgroups. *Alzheimer's & Dementia*, *7*(3), 263–269. https://doi.org/10.1016/j.jalz.2011.03.005

Mitchell, A. J., & Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia-meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica*, *119*(4), 252–265. https://doi.org/10.1111/j.1600-0447.2008.01326.x

Moghekar, A., Li, S., Lu, Y., Li, M., Wang, M. C., Albert, M., & O'brien, R. (2013). CSF biomarker changes precede symptom onset of mild cognitive impairment. *Neurology*, *81*(20), 1753–1758. https://doi.org/10.1212/01.wnl.0000435558.98447.17

Morris, J. C. (1993). The clinical dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–2414. https://doi.org/10.1212/WNL.43.11.2412-a

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: A concept in evolution. *Journal of Internal Medicine*, *275*(3), 214–228. https://doi.org/10.1111/joim.12190

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (2001). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(3), 303–308. https://doi.org/10.1001/archneur.56.3.303

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Jr., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology*, *37*(3), 323–329. https://doi.org/10.1093/geronj/37.3.323

Reitan, R. M. (1958). Validity of the Trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*(3), 271–276. https://doi.org/10.2466/pms.1958.8.3.271

Rey, A. (1964). *L'examen clinique en psychologie*. Presses Universitaires de France.

Russo, M. J., Bueri, J., Alba-Ferrara, L., Labos, E., Sevlever, G. E., & Allegri, R. F. (2026). Recognition memory in individuals with mild cognitive impairment and Alzheimer's dementia: A systematic review. *Applied Neuropsychology Adult*, 1–21. https://doi.org/10.1080/23279095.2025.2610375

Russo, M. J., Campos, J., Martín, M. E., Clarens, M. F., Sabe, L., & Allegri, R. F. (2013). Índice de discriminabilidad en memoria de reconocimiento en deterioro cognitivo leve amnésico y demencia tipo Alzheimer leve: Un estudio preliminar. *Neurología Argentina*, *5*(4), 241–249. https://doi.org/10.1016/j.neuarg.2013.08.002

Russo, M. J., Campos, J., Vázquez, S., Sevlever, G., & Allegri, R. F. (2017). Adding recognition discriminability index to the delayed recall is useful to predict conversion from mild cognitive impairment to Alzheimer's disease in the Alzheimer's disease neuroimaging initiative. *Frontiers in Aging Neuroscience*, *9*, 46. https://doi.org/10.3389/fnagi.2017.00046

Russo, M. J., Cohen, G., Campos, J., Martín, M. E., Clarens, M. F., Sabe, L., Allegri, R. F., & Allegri, R. F. (2016). Usefulness of discriminability and response bias

indices for the evaluation of recognition memory in mild cognitive impairment and Alzheimer disease. *Dementia and Geriatric Cognitive Disorders*, *41*(1–2), 1–14. https://doi.org/10.1159/000452255

Russo, M. J., Cohen, G., Méndez, P. C., Campos, J., Nahas, F. E., Surace, E. I., Sevlever, G., Gustafson, D., Guinjoan, S., Allegri, R., & Sevlever, G. (2016). Predicting episodic memory performance using different biomarkers: Results from argentina-alzheimer's disease neuroimaging initiative. *Neuropsychiatric Disease and Treatment*, *12*, 2199–2206. https://doi.org/10.2147/NDT.S107051

Schmidt, M. (1996). *Rey Auditory verbal Learning test: A handbook*. Western Psychological Services.

Simons, J. S., & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, *4*(8), 637–648. https://doi.org/10.1038/nrn1178

Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P., & Barnes, C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Reviews Neuroscience*, *12*(10), 585–601. https://doi.org/10.1038/nrn3085

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology General*, *117* (1), 34–50. https://doi.org/10.1037/0096-3445.117.1.34

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods Instruments & Computers*, *31*(1), 137–149. https://doi.org/10.3758/BF03207704

Stark, S. M., Yassa, M. A., Lacy, J. W., & Stark, C. E. (2015). A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia*, *51*(12), 2442–2449. https://doi.org/10.1016/j.neuropsychologia.2012.12.014

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford University Press.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised*. Psychological Corporation.

Weintraub, S., Salmon, D., Mercaldo, N., Ferris, S., Graff-Radford, N. R., Chui, H., Morris, J. C., DeCarli, C., Foster, N. L., Galasko, D., Peskind, E., Dietrich, W., Beekly, D. L., Kukull, W. A., & Morris, J. C. (2009). The Alzheimer's disease Centers' uniform data Set (UDS): The neuropsychologic test battery. *Alzheimer Disease & Associated Disorders*, *23*(2), 91–101. https://doi.org/10.1097/WAD.0b013e318191c7dd

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152–176. https://doi.org/10.1037/0033-295X.114.1.152

Wolk, D. A., Dunfee, K. L., Dickerson, B. C., Aizenstein, H. J., & DeKosky, S. T. (2011). A medial temporal lobe division of labor: Insights from memory in aging and early Alzheimer disease. *Hippocampus*, *21*(5), 461–466. https://doi.org/10.1002/hipo.20779

Wolk, D. A., Signoff, E. D., & DeKosky, S. T. (2008). Recollection and familiarity in amnestic mild cognitive impairment: A global decline in recognition memory. *Neuropsychologia*, *46*(7), 1965–1978. https://doi.org/10.1016/j.neuropsychologia.2008.01.017

Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*(1), 37–49. https://doi.org/10.1016/0022-3956(82)90033-4

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*(3), 441–517. https://doi.org/10.1006/jmla.2002.2864