



**HAL**  
open science

# The Memory Binding Test Detects Early Subtle Episodic Memory Decline in Preclinical Alzheimer's Disease: A Longitudinal Study

Filipa Rapos Pereira, Nathalie George, Gianfranco Dalla Barba, Bruno Dubois, Valentina La Corte

► **To cite this version:**

Filipa Rapos Pereira, Nathalie George, Gianfranco Dalla Barba, Bruno Dubois, Valentina La Corte. The Memory Binding Test Detects Early Subtle Episodic Memory Decline in Preclinical Alzheimer's Disease: A Longitudinal Study. *Journal of Alzheimer's Disease*, 2024, 98 (2), pp.465-479. 10.3233/JAD-230921 . hal-04519805

**HAL Id: hal-04519805**

**<https://hal.science/hal-04519805v1>**

Submitted on 25 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Memory Binding Test Detects Early Subtle Episodic Memory Decline in Preclinical Alzheimer's Disease: A Longitudinal Study

Filipa Raposo Pereira<sup>a,b,\*</sup>, Nathalie George<sup>a</sup>, Gianfranco Dalla Barba<sup>c</sup>, Bruno Dubois<sup>a,b,d</sup>,  
Valentina La Corte<sup>b,e,f</sup> and the INSIGHT-preAD study group

<sup>a</sup>*Institut du Cerveau – Paris Brain Institute – ICM, INSERM, U 1127, CNRS, UMR 7225' APHP, CENIR, Centre MEG-EEG, Sorbonne Université, Hôpital de la Pitié-Salpêtrière, Paris, France*

<sup>b</sup>*Department of Neurology, Institute of Memory and Alzheimer's Disease (IM2A), Pitié-Salpêtrière University Hospital, Assistance Publique – Hôpitaux de Paris (AP-HP), Paris, France*

<sup>c</sup>*Dipartimento di Scienze della Vita, Università degli Studi di Trieste, Trieste, Italy*

<sup>d</sup>*Department of Neurology, Centre of Excellence of Neurodegenerative Disease (CoEN), ICM, CIC Neurosciences, Assistance Publique – Hôpitaux de Paris (AP-HP), Hôpital Pitié-Salpêtrière, Sorbonne Université, Paris, France*

<sup>e</sup>*Laboratoire Mémoire Cerveau et Cognition (UR 7536), Institut de Psychologie, Université Paris Cité, Boulogne-Billancourt, France*

<sup>f</sup>*Institut Universitaire de France, Paris, France*

Accepted 15 January 2024

Pre-press 22 February 2024

## Abstract.

**Background:** The asymptomatic at-risk phase might be the optimal time-window to establish clinically meaningful endpoints in Alzheimer's disease (AD).

**Objective:** We investigated whether, compared with the Free and Cued Selective Reminding Test (FCSRT), the Memory Binding Test (MBT) can anticipate the diagnosis of emergent subtle episodic memory (EM) deficits to an at-risk phase.

**Methods:** Five-year longitudinal FCSRT and MBT scores from 45 individuals matched for age, education, and gender, were divided into 3 groups of 15 subjects: A $\beta$ -/controls, A $\beta$ +/stable, and A $\beta$ +/progressors (preclinical-AD). The MBT adds an associative memory component (binding), particularly sensitive to subtle EM decline.

**Results:** In the MBT, EM decline started in the A $\beta$ +/progressors (preclinical-AD) up to 4 years prior to diagnosis in delayed free recall (FR), followed by decline in binding-associated scores 1 year later. Conversely, in the FCSRT, EM-decline began later, up to 3 years prior to diagnosis, in the same group on both immediate and delayed versions of FR, while on total recall (TR) and intrusions decline started only 1 year prior to diagnosis.

**Conclusions:** The MBT seems more sensitive than the FCSRT for early EM-decline detection, regarding the year of diagnosis and the number of scores showing AD-linked EM deficits (associated with the AD-characteristic amnesic hippocampal syndrome). Considering the MBT as a detection tool of early subtle EM-decline in an asymptomatic at-risk phase, and the

---

\*Correspondence to: Filipa Raposo Pereira, Paris Brain Institute, 47 Boulevard de l'Hôpital, Hôpital Pitié-Salpêtrière, Paris, France. E-mail: filipa.rapospereira@icm-institute.org.

FCSRT as a classification tool of stages of EM-decline from a preclinical phase, these tests ought to potentially become complementary diagnostic tools that can foster therapies to delay cognitive decline.

Clinical trial registration title: Electrophysiological markers of the progression to clinical Alzheimer disease in asymptomatic at-risk individuals: a longitudinal event-related potential study of episodic memory in the INSIGHT pre-AD cohort (acronym: ePARAD).

Keywords: Alzheimer's disease, amyloid- $\beta$ , at-risk, episodic memory, free and cued selective reminding test, memory binding test, preclinical Alzheimer's disease

## INTRODUCTION

Alzheimer's disease (AD) remains the most disabling dementia and the greatest public health problem in aging populations [1]. At present AD has no cure, and the way diagnosis is performed imposes major obstacles to the early application of promising therapies, that can at best delay cognitive hallmarks of AD such as episodic memory (EM) decline (i.e., deficits in the recollection of personal events and their spatiotemporal context) [2–6].

In typical AD, reported subjective memory complaints (SMC) tend to be the earliest signs of EM deficits ( $\pm 4$ –8 years prior to other executive failures), which are AD's most prominent cognitive dysfunction, and highly predictive of incident dementia [2, 5, 7–11]. AD assessments largely result from SMC, often self-reported when the patient is already on the path to clinical symptomatology [2, 5, 7–9, 11]. This prevents an early onset detection of EM decline [2, 5, 7–9, 11] and delays AD-diagnosis to late preclinical stages when biomarkers such as amyloid- $\beta$  burden ( $A\beta+$ ) or neurodegeneration are confirmed [2, 4–6, 8, 9]. The anticipation of AD-diagnosis to an asymptomatic at-risk stage (i.e., when individuals are within normal cognitive capacity according to normative cut-offs and biomarkers may or may not be present), is further challenged by the typical presence of  $A\beta$  burden, neurodegeneration, and SMC among aging populations [2, 9, 12–15]. Nevertheless, growing evidence suggests that this might be the ideal time-window to apply effective therapeutics that delay cognitive decline, which has contributed to the shift in the focus of AD research to earlier preclinical stages [3, 14, 16, 17]. However, further research in this direction is needed.

EM decline is the first cognitive deficit to emerge in typical AD and therefore naturally emerges as a promising cognitive marker [3, 6, 9]. Deterioration in EM processing starts with a poor capacity to encode information to be remembered, mostly associated with problems in executive learning strategies,

evolving to storage impairment linked to inefficient retrieval, characteristic of the amnesic hippocampal syndrome (AHS), a clinical phenotype of AD [5–8, 11, 18]. EM decline can be particularly significant when semantic material is used, since evidence suggests that partial degradation of semantic networks may parallel initial EM degradation [5, 11, 19–21]. Therefore, in AD, EM is mostly assessed with the Free and Cued Selective Reminding Test (FCSRT) [2, 3, 8, 10, 14, 17, 18, 22]. The test employs controlled learning by pairing a word to be remembered with a semantic cue used at the encoding and recall stages to maximize retrieval, which ensures proper encoding and minimizes the use of individual learning strategies [5, 10, 18, 21, 23]. EM decline is based exclusively on the FCSRT scores that integrate the official Preclinical Alzheimer's Cognitive Composite (PACC) [2, 10, 17]. These scores are: free recall (FR), signaling retrieval (i.e., accessibility), and total recall (TR; with aiding cues), signaling storage (i.e., availability) [2, 5, 6, 10, 18, 23–25]. FR decline underlies encoding deficits mostly linked to executive dysfunction, which tends to be the first score to deteriorate (6.6–7.3 years), being often related to  $A\beta$  burden [5, 6, 10, 23, 25]. TR decline mostly underlies storage deficits that are typical markers of AD-AHS, appearing at late preclinical stages, when semantic cues start failing to overcome retrieval deficits [5, 7, 10, 11, 18, 22, 24, 26]. However, despite its diagnostic value at a preclinical stage, currently the FCSRT presents three main limitations to anticipate diagnosis to an at-risk phase: 1) the FR/TR decline rates vary significantly inter-individually, and although a 'Stages of Objective Memory Complaints' (SOMI) system has been recently proposed to classify EM-decline accounting for these variability, the use of SOMI is still restricted to preclinical and clinical stages; 2) only TR deficits are characteristic of AD-AHS hence, when retrieval impairment (i.e., FR) appears in isolation, it might result from executive or attention deficits that are not necessarily AD-related; 3) by using one list of words, FCSRT-TR is prone to the ceiling effect, restricting

comparison of its scores to cognitive aging equivalents, and delaying detection of TR alterations [2, 5, 7, 10, 16, 25, 27].

The Memory Binding Test (MBT) avoids these limitations by using two lists of words with semantic cues as ‘unifying units’, binding 1 item of list A with only 1 item of list B [16, 19, 27]. This introduces the associative memory component, or binding, (i.e., the ability to encode and recall distinct information as part of a complex unit), sourced mainly in the hippocampus [16, 19, 27]. Accordingly, beyond FR and TR scores similar to those of the FCSRT but accounting for double the items (i.e., each pair of words belongs to a semantic category given as a cue for the TR), the MBT introduces two binding-specific measures (accounting for the total number of items [TIP] or the number of pairs of items [PIP]), suggested to be more sensitive to subtle EM alterations [16, 19, 27]. The use of two lists allows the comparison of TR and binding scores with either an individual’s own performance, their initial capacity, or AD-patient’s performance, eliminating the ‘ceiling levels’ observed in the variability of high TR scores, which are not detected with the FCSRT [16, 19, 25, 27–29]. In AD, MBT deficits have been linked to AD biomarkers and an increased risk of AD-related amnesia and dementia [10, 14, 16, 19, 25, 28–30]. Globally, research suggests that the binding of semantic information increases the sensitivity to subtle EM alterations emerging at an at-risk phase, while individuals remain within normal psychometric capacity according to the FCSRT [10, 14, 16, 19, 21, 25, 28–30].

Overall, the MBT is practical and easy to administer, with a good test-retest reliability, that avoids arbitrary cut-offs and local normative samples, adding considerable benefits in the triage for earlier AD treatments, initial prevention trials, and imaging testing [2, 3, 16, 19, 29, 30]. Currently, the FCSRT validity, the official tool for diagnosis of AD-linked EM decline, is based on multiple studies assessing its normative and psychometric properties [5, 10, 17, 21, 23, 25]. However, despite remaining highly relevant for characterizing stages of EM decline, it is able to detect these deficits only at late preclinical stages [5, 10, 16, 23]. Alternatively, the MBT is suggested to be more sensitive to subtle EM decline when this starts appearing at an early preclinical stage, showing great potential to become a complementary diagnostic tool [16, 19, 29].

To test this, we had two main objectives: 1) Retrospectively assess the performance of the MBT

main scores (FR, TR, and binding) during a 5-year follow-up period, in a sample of cognitively normal individuals at-risk for AD (>6months of subjective memory complaints) divided in 3 groups, namely controls without A $\beta$  burden, individuals with A $\beta$  burden but stable (without cognitive decline), and individuals who progressed to prodromal AD (considered only before their diagnosis: preclinical AD); 2) In the same sample, during the same follow-up period, assess the MBT and FCSRT performance in their main parallel scores (FR and TR), with the aim of comparing the initial session at which significant decline in EM starts occurring and whether decline in binding is detected earlier.

## METHODS

### *Participants*

The 318 elderly participants included in the multidisciplinary INSIGHT-pre AD cohort were recruited from the Institute for Memory and Alzheimer’s disease (IM2A) at the Pitié-Salpêtrière Hospital of Paris (monocentric) and followed-up for 5 years. Inclusion criteria were: 70–85 years of age;  $\geq 6$  months of self-reported SMC; unimpaired cognition (the Mini-Mental State Examination [MMSE] total score  $\geq 27/30$ , and the FCSRT-TR score  $\geq 41/48$ ); null clinical dementia rating (score = 0); good visual and auditory acuity [18, 31, 32]. Exclusion criteria were the following: being under guardianship or in a nursing facility; diagnosis of AD or any other neurological disease; presented illiteracy; any impossibility to perform a brain MRI. AD diagnosis was assessed by an independent committee of two neurologists, a neuropsychologist, and a neuro imaging expert in the event of an individual’s cognitive decline in two consecutive neuropsychological assessments presenting A $\beta$  burden (A $\beta$ -SUVR  $> 0.7918$ ) and persistent AHS [33]. Therefore, ‘progression to prodromal AD’ was defined as an early symptomatic phase of AD characterized by the specific clinical phenotype consisting on the presence of AHS and positive pathophysiological biomarkers. AD-diagnosis meant removal from the INSIGHT-preAD cohort. During follow-up 15 subjects progressed to AD and therefore, as a preliminary study, we decided to carefully match them in terms of age, sex, education level, and A $\beta$  burden and neurodegeneration SUVR, with two groups of 15 subjects each. Hence, here we present a subset of 45 cognitively normal elderly individuals retrospectively divided in three groups: 15 A $\beta$  burden negative

participants ( $A\beta$ -/controls); 15  $A\beta$  burden positive non-progressors to prodromal AD ( $A\beta$ +/stable); and 15  $A\beta$  burden positive converters to prodromal AD ( $A\beta$ +/progressors [preclinical-AD]). Based on the higher risk for AD introduced by the presence of  $A\beta$  burden, we considered  $A\beta$ - participants as controls, and  $A\beta$ +/stable participants at a higher risk of progressing to AD.

### *Ethics*

All aspects of this study were designed in full compliance with the French law n° 2004-806 (9 August 2004), the Good Clinical Practice principles (I.C.H version 4 of 1 May 1996 and the decision of 24 November 2006), the guidelines of the Helsinki Declaration (Ethical Principles for Medical Research involving Human Subjects, Tokyo 2004), and the human ERP research technical guidelines. The ethics committee of the Pitie-Salpêtrière University Hospital fully approved the INSIGHT-PreAD protocol and the INSIGHT-preAD scientific committee approved this project.

### *Consent statement*

After the study conditions had been presented, a signed informed consent and willingness to commit to the longevity of the project was required from each participant. Drop-out was possible at any time if participants freely wished. Participation in the study automatically ceased in the event of diagnosis with another condition.

### *Procedure*

The INSIGHT-pre AD initiative is an extensively multidisciplinary observational longitudinal study (5 years). It combines demographic, neuropsychological, neurobiological, genetic, neuroimaging, and electrophysiological data collected mono centrally to investigate the preclinical progression of cognitively normal aging individuals aiming at developing comprehensive individual characterization during the at-risk phase (SMC > 6 months) to identify markers that can anticipate diagnosis [3]. The subset of data presented here comprises the following: positron emission tomography (PET) imaging scans collected at baseline, to assess  $A\beta$  burden and neurodegeneration; neuropsychological evaluation collected at baseline and at every 12 months (to assess general cognitive efficiency, EM capacity, and executive

functioning) [18, 31]. In all sessions the neuropsychological assessments were performed in the morning and the neuroimaging assessments administered in the afternoon of the day of the experiment, with breaks introduced between the assessments throughout the day.

### *PET acquisitions*

The regional standard uptake value ratio (SUVr) of either  $A\beta$  burden or neurodegeneration was collected with the Philips Gemini GXL CR-PET scanner.

$A\beta$  burden SUVr was measured 50 min after an injection of 370 MBq (10mCi)  $^{18}\text{F}$ -florbetapir (marker of neocortical deposition of  $A\beta$  plaques). Regional values were collected from the following regions of interest (ROIs): left and right precuneus, posterior cingulum, anterior cingulum, associative parietal cortex, associative temporal cortex, orbitofrontal cortex, and whole cerebellum plus pons (reference region).  $A\beta$  positive status ( $A\beta$ +) was based on the SUVr > 0.7918 cut-off defined by a linear conversion of the CAEN method averaging the mean SUVr of the ROIs to the reference region previously performed for the INSIGHT-preAD study [33].

Neurodegeneration was considered in this study as an expression of hypometabolism, or the deficient uptake of cortical glucose. FDG SUVr was measured after 30 minutes of an injection of 2 MBq/kg  $^{18}\text{F}$ -FDG (measuring cortical glucose uptake in the following ROIs: posterior cingulate cortex, inferior parietal lobe, precuneus, inferior temporal gyrus, and pons (reference region). Neurodegeneration positive (N+) status was based on SUVr < 2.27 cut-off obtained from a similar linear method described above averaging the mean SUVr of the ROIs to the reference region created by Jack et al. [4].

### *Neuropsychological assessments*

The French socio-cultural (NSC) scale was used to assess education scores from levels 1 (illiterate level) to 8 (2 years of higher education after a bachelor's degree) as a proxy for intelligence. General cognitive efficiency was assessed with the 30-point MMSE, where a score < 27 is suggestive of cognitive impairment and/or dementia, requiring a more detailed cognitive examination [3, 31]. Executive function was assessed with the Frontal Assessment Battery (FAB) to discriminate AD from frontal dementias [34]. A score  $\leq 12$  out of 18 is considered indica-

tive of AD [34]. The 15-item French version of the McNair Frequency of Forgetting Questionnaire (McNair Questionnaire) was used to assess the severity and frequency of SMC when reported by the participant (from M0) and by a companion (from M12) [35]. A score  $\geq 15$  is suggestive of abnormal SMC levels [35].

#### *Free and Cued Selective Reminding Test (FCSRT)*

The FCSRT begins with the consecutive presentation of 4 cards, each containing 4 different words to be remembered, amounting to 16 in total [6, 18, 24, 36]. On each card, participants are asked orally to identify the item (e.g., grapes) that corresponds to a unique category cue (e.g., fruit; learning and encoding phase). This is followed by three immediate FR trials in which is required to freely recall as many items as possible without the aid of semantic cues. Immediately after each FR trial, follows a cue recall (CR) trial, in which semantic cues are presented for the items not remembered until all missing items have been retrieved. The test includes three TR (FR plus CR) trials intercalated with interference tasks of 20 s of backwards counting. Twenty minutes later delayed FR (DFR) and delayed CR (DCR) trials are administered for the same semantic context. The FR (sum of all immediate FR trials, range 0–48), a measure of ‘availability’ of information, and TR (the sum of 3FR and 3CR trials, range 0–48), a measure of ‘accessibility’ to stored information, suggestive of temporal-limbic amnesia when  $\leq 41$ , are the main FCSRT scores used to assess the two different types of AD-recall deficits [3, 18, 23, 30].

#### *Memory Binding Test (MBT)*

The MBT (formerly, memory capacity test) consists in the learning of two alternate lists of 16 words presented one after the other in a similar manner to the FCSRT [16, 27]. The two lists are linked through 16 identical semantic cues (i.e., semantic cue = insect, word List A = flea, word List B = ant) that function as unifying units to promote binding, proxy to associative memory assessment. The test starts by learning List A (encoding) immediately followed by a CR trial (CRa; 0–16 score), in which semantic cues are presented by the examiner (duration 5 s). The participant is asked to verbally identify the corresponding item, to ensure controlled learning

and encoding specificity. The same procedure is then repeated for List B and semantic cues presented in the same order (CRb; 0–16 score). It follows a total CR ( $TCR_{A+B}$ ; paired condition), in which semantic cues are provided orally by the examiner in the same order. Participants are asked to identify both items from the two lists in any order (within 10 s each), originating two measures. The total number of items recalled in the paired condition (TIP; 0–32 score), quantifying all items correctly associated to each cue (paired or unpaired), and the number of pairs cued recall in the paired condition (PIP; 0–16 score) quantifying only the correct pairs for each cue. It follows a FR trial in which participants are required to recall as many items as possible from the two lists in no particular order and without the aid of cues (0–32 score). Thirty minutes after it follows a DFR (0–32 score) and a delayed TIP (DTIP; 0–32 score). Immediately after  $TCR_{A+B}$ , participants are asked to identify the origin (List A or B) of each word constituting a word-pair with the question “Which word was in list A and which word was in list B?”. This constitutes the immediate and delayed source memory recall (SMR and DSMR, respectively) scores, underlying retrieval of the encoding context are then quantified in percentage. Different modalities of intrusions are also recorded: total number of intrusions (T\_int); semantically related intrusions (extra-list intrusions); semantically unrelated intrusions (extra-category intrusions); and intrusions introduced by words originating from List A when List B was required to be recalled (prior list of intrusions).

#### *Statistical analysis*

The statistical software R studio (version 4.2.1; <https://www.r-project.org/>) was used to assess the neuropsychological longitudinal data collected for 5 years at every 12 months from baseline (M0; M12; M24; M36; M48; M60).

#### *Demographic and clinical characterization of the groups at baseline (M0)*

Demographic and clinical traits were examined at baseline (M0) to assess potential differences between our groups of interest ( $A\beta^-$ /controls,  $A\beta^+$ /stable,  $A\beta^+$ /progressors). One-way Welch analyses of variance (ANOVA; car R package, version 3.1.0), considering all the groups as a between-subject factor (to ensure homogeneity between them), were used to assess age, sex, education level and SMC. Differences

in A $\beta$  SUV<sub>r</sub> were tested between the A $\beta$ +/*stable* and the A $\beta$ +/*progressors* (preclinical-AD). Significance was considered when  $p < 0.05$ .

The neuropsychological tests considered as inclusion criteria in the study (i.e., MMSE, FCSRT\_TR, FAB) were assessed at baseline with general linear models (Glm; 'stats' R package, version 4.2.1) to control for potential confounders known to influence cognitive performance introduced by differences in age, sex, and education level. In these models the groups were considered as a fixed between-subject factor and age, sex, and education level considered as fixed-effect covariates of no interest. Age was mean centered at baseline and considered as a numerical variable, while sex and education level were introduced as categorical variables. Games-Howell tests (rstatix R package, version 0.7.0) were used to assess post-hoc comparisons and results were corrected for multiple comparisons.

#### *FCSRT and MBT analysis*

We used a statistical model considering the three groups described above to assess our fixed-effects of interest as follows: A $\beta$ -/*controls* versus A $\beta$ +/*stable* contrast, to isolate A $\beta$  burden effect; A $\beta$ +/*stable* versus the A $\beta$ +/*progressors* (preclinical-AD) contrast to isolate the progression to prodromal AD effect. Linear mixed-effects (lmer) models (lme4R package, version 1.1.29) were used to analyze the outcome variables for their ability to account for missing responses and assess longitudinal data. Recording 'session' (6 levels: M0, M12, M24, M36, M48, M60) was considered as a fixed-effect repeated measures experimental factor, and '1| participant iD' was considered as a random-effect factor with a random intercept per subject. Variability introduced by differences in age, sex, and education level was minimized by considering these variables as covariates of no interest. For the outcome measures in which main effects were found to be significant, post-hoc analysis was performed using emmeans (R package, version 1.7.5) between the groups underlying our effects of interest (i.e., A $\beta$  burden or progression to prodromal AD) and controlled for multiple comparisons. Statistical significance was achieved when  $p < 0.05$ .

#### *Additional control analyses*

As a complement to the characterization of the groups, body mass index (BMI [weight/height<sup>2</sup>]; because of its link with increased glucose) and licit substance use of alcohol and cigarettes, were assessed

through a 1-way Welch ANOVA at baseline in the same manner as the assessment of the demographic variables described above.

Longitudinal differences in SMC between the groups, self-reported by participants or reported by their companions from M12 onwards (since this was the first session recorded with the companions) were also assessed with 1-Way Welch ANOVAs. These models were similar to the ones described at baseline. However, session was added as a fixed within-subject (i.e., repeated-measure) factor, and the interaction between group\_status\*session was considered. Finally, similar lmer models to the ones used on the FCSRT and the MBT analysis were applied to the MMSE and FAB total scores to also assess their longitudinal evolution in the three groups of interest.

We report  $F$  and  $p$ -values of significant effects for the 1-way Welch ANOVA, Glm, and lmer analysis. Multiple comparisons correction was performed using a multivariate  $t$  distribution correction method (emmeans R package, version 1.7.5).

## RESULTS

### *Demographic and clinical characterization of the groups at baseline (M0)*

From the 318 elder participants with SMC at-risk for AD initially recruited in the INSIGHT-preAD cohort, we retrospectively selected a subsample of 45 individuals matched for age, sex, education level, and A $\beta$  and neurodegeneration SUV<sub>r</sub>. This subset consisted of 15 A $\beta$ -/*controls* and 15 A $\beta$ +/*stable*, based on the 15 A $\beta$ +/*progressors* (preclinical-AD) who progressed to prodromal AD during the 5-year follow-up period (evaluated here only until their diagnosis; Supplementary Material). At baseline the population was on average  $77.3 \pm 1.5$  years old [76–79 years], with a slight majority of women 56% ( $n = 25$ ), and with a high education level  $6.4 \pm 2.1$  (bachelor undergraduate level; Table 1). No significant differences in these variables were found between the groups at baseline. As expected, significant group differences were found in A $\beta$  SUV<sub>r</sub> [ $F(1,27) = 8.8, p = 0.006$ ] between the groups A $\beta$ +, with post-hoc tests evidencing a significantly higher A $\beta$  burden in the A $\beta$ +/*progressors* (preclinical-AD) than in the A $\beta$ +/*stable* ( $p = 0.006$ ; Table 1).

The overall mean for each neuropsychological score required as an inclusion criterion for the INSIGHT-preAD cohort at entry level (M0) was met by all the groups: MMSE\_total score  $28.4 \pm 0.8$

Table 1  
Demographic and clinical characterization of the three groups of interest: Aβ-/controls, Aβ+/stable, and Aβ+/progressors (preclinical-AD)

score_name	M <sub>0</sub>			M <sub>60</sub>			group-status (M0)	group-status*
	Controls	Stable	Progressors	Controls	Stable	Progressors		
	Aβ- (n = 15) mean ± SD	Aβ+ (n = 15) mean ± SD	Aβ+ (n = 15) mean ± SD	Aβ- (n = 14) mean ± SD	Aβ+ (n = 12) mean ± SD	Aβ+ (n = 8) mean ± SD	p	p
Age	77.5 ± 3.4	77.2 ± 3.0	77.4 ± 3.5	–	–	–	0.97 <sup>a</sup>	–
Gender (%)	M = 7(47) F = 8(53)	M = 6(40) F = 9(60)	M = 7(47) F = 8(53)	–	–	–	0.92 <sup>a</sup>	–
Education level	6.4 ± 2.2	6.5 ± 2.0	6.4 ± 2.3	–	–	–	0.98 <sup>a</sup>	–
Aβ SUVR [Aβ+ >0.7918]	0.7 ± 0.1	1.0 ± 0.2	1.2 ± 0.2	–	–	–	0.006 <sup>**,a</sup>	–
BMI [>25 = overweight]	24.5 ± 3.9	25.7 ± 3.7	25.3 ± 3.3	–	–	–	0.69 <sup>a</sup>	–
SMC_companion (n) [normal < 15; baseline = M12]	9.6 ± 7.9 (14)	6.3 ± 5.2 (12)	10.0 ± 7.7 (13)	11.3 ± 8.6 (12)	11.9 ± 10.6 (10)	20.8 ± 10.4 (6)	–	0.07 <sup>b</sup>
SMC_participant [normal < 15]	14.5 ± 6.5	10.6 ± 4.0	11.8 ± 5.95	16.5 ± 8.1	14.0 ± 5.1	16.9 ± 6.7	–	0.29 <sup>b</sup>
MMSE (min-max) [normal ≥ 27/30]	28.7 ± 1.0 [27–30]	28.4 ± 0.7 [27–30]	28.2 ± 0.7 [27–29]	28.6 ± 1.7 [25–30]	28.3 ± 1.2 [26–30]	26.5 ± 1.9 [24–29]	–	0.61 <sup>c</sup>
FAB (min-max) [normal ≥ 12/18]	17.0 ± 1.2 [14–18]	16.3 ± 1.7 [12–18]	16.0 ± 2.1 [12–18]	17.0 ± 0.9 [15–18]	15.7 ± 2.1 [12–18]	14.5 ± 1.6 [12–17]	–	0.11 <sup>c</sup>
FCSRT_TR (min-max) [normal ≥ 41/48]	45.9 ± 2.2 [41–48]	46.1 ± 1.9 [42–48]	44.7 ± 2.5 [41–48]	46.6 ± 2.3 [41–48]	46.3 ± 2.0 [43–48]	34.9 ± 7.1 [21–41]	–	<0.001 <sup>***,c</sup>

SD, standard deviation; Aβ SUVR, Amyloid Beta Standardized Uptake Value Ratio; MMSE, Mini-Mental State Examination; FCSRT, Free and Cued Selective Reminding Test; FAB, Frontal Assessment Battery. The table represents the mean and standard deviation of each group at the first and last sessions (M0, M60, respectively). Significance is reported for each outcome variable based on the following p-values: \*\*\* $p \leq 0.001$ ; \*\* $p \leq 0.01$ ; \* $p \leq 0.05$ . <sup>a</sup>Baseline (M0) Welch Analysis of variance (ANOVA). <sup>b</sup>Longitudinal Welch ANOVA. <sup>c</sup>Longitudinal General linear model (Glm) analysis.



( $\geq 27/30$ ), the FAB total score  $16.4 \pm 1.7$  ( $\geq 12/18$ ), and the FCSRT TR  $45.7 \pm 2.2$  ( $\geq 41/48$ ). All participants were considered at-risk for AD based on their self-report of SMC at consultation for at least 6 months (McNair Questionnaire). At baseline all the groups showed normal SMC levels ( $< 15$ ). And no significant were observed between them. However, at baseline, the A $\beta$ -/controls self-reported the highest score of SMC, approximating an abnormal level of 14.5 at entry level.

### FCSRT analysis

Longitudinally, a significant group\_status\*session interaction was found for the main score of the FCSRT, which were: FR [F(10,179)=7.3,  $p < 0.001$ ], DFR [F(10,182)=10.5,  $p < 0.001$ ], TR [sum of cued and FR; F(10,177)=8.3,  $p < 0.001$ ], DTR [F(10,179)=8.8,  $p < 0.001$ ], and intrusions [F(10,178)=4.0,  $p < 0.001$ ; Table 2].

Post-hoc analysis showed that the A $\beta$ +/progressors (preclinical-AD) scored worse than the A $\beta$ -/controls and the A $\beta$ +/stable from session M24 onwards on the FR ( $p = 0.0008$ ,  $p = 0.024$ , respectively) and DFR ( $p = 0.0001$ ,  $p = 0.04$ , respectively), and only from session M48 onwards on the TR ( $p = 0.0002$ ,  $p = 0.0009$ , respectively), DTR ( $p < 0.0001$ ,  $p < 0.0001$ , respectively), and intrusions ( $p = 0.003$ ,  $p = 0.0001$ , respectively; Fig. 1).

No effect of A $\beta$  burden per se was observed on the FCSRT.

### MBT analysis

A significant group\_status\*session interaction was found on every main score of the MBT related to the paired condition (list A+B), namely: FR [F(10,177)=2.9,  $p = 0.002$ ]; DFR [F(10,163)=2.9,  $p = 0.002$ ]; TCR<sub>A+B</sub> [F(10,175)=4.8,  $p < 0.001$ ]; PIP [F(10,176)=3.2,  $p < 0.001$ ]; TIP [F(10,160)=4.3,  $p < 0.001$ ]; DTIP [F(10,160)=5.2,  $p < 0.001$ ]; SMR [F(10,162)=2.7,  $p = 0.005$ ]; DSMR [F(10,165)=1.8,  $p = 0.06$ ], as well as on the total number of intrusions [T\_int; F(10,176)=3.6,  $p < 0.001$ ] and on the extra category of intrusions [extra\_int; F(10,176)=3.8,  $p < 0.001$ ; Table 3].

Post-hoc analysis showed that right from baseline (up to 5 years prior to diagnosis), there was a tendency for the A $\beta$ +/progressors (preclinical-AD) to score worse than the A $\beta$ +/stable on the TCR<sub>A+B</sub> and PIP ( $p = 0.07$ ,  $p = 0.06$ , respectively). However, scores started to be consistently and significantly worse

among the A $\beta$ +/progressors (preclinical-AD) when compared with the A $\beta$ -/controls and the A $\beta$ +/stable only from session M12 onwards on DFR ( $p = 0.005$ ,  $p = 0.03$ , respectively). From session M24 this pattern was extended to the FR, TCR<sub>A+B</sub>, PIP, TIP, and DTIP, all scores linked to the paired condition of the test associated with binding. It was not until session M60 that these differences were also verified on the SMR, DSMR, T\_int, and extra\_int scores (Fig. 2).

No effect of A $\beta$  burden per se was found on any MBT score.

### Additional control analyses

All groups showed a normal use of alcohol and cigarettes, and no significant differences were found between them. No BMI differences were found between the groups. However, the BMI of the A $\beta$ +/stable and A $\beta$ +/progressors (preclinical-AD) groups was high (25.6 and 25.3, respectively) and could be considered to be in the lower part of the official window for overweight classification (25–29.9). Since SMC were recorded from session M0 on the participants and only from session M12 on the companions, we considered the longitudinal analysis of SMC from M12 onwards. In the SMC report of the companions a tendency for differences were found between the groups [F(14,58)=1.7,  $p = 0.07$ ]. At session M12, all the companions reported normal scores of SMC. Self-reported, the SMC from the A $\beta$ +/progressors (preclinical-AD) reached abnormal levels only from session M36 onwards (16.5), but no significant differences were found between the groups. Interestingly, the A $\beta$ -/controls self-reported more SMC than the other two groups from session M12-36 and were borderline abnormal (normal  $< 15$ ) until session M48, reaching a pathological clinical level at session M60. While among the A $\beta$ +/progressors (preclinical-AD), SMC became pathological from M48.

No longitudinal differences were observed between the groups in the MMSE and FAB scores.

## DISCUSSION

In line with our hypothesis, the MBT was highly sensitive in detecting EM decline in multiple scores characteristic of the AD amnesic hippocampal syndrome (AHS) already at an early preclinical stage. Moreover, the MBT showed AD-specific EM decline in a higher number of scores than the FCSRT, and two sessions earlier. Hence, we suggest the potential

Table 2  
The Free and Cued Selective Reminding Test (FCSRT) main scores at baseline and after 5 years of follow-up

score_name	M <sub>0</sub>			M <sub>60</sub>			F <sub>df1,df2</sub> = value	group_status*sess <i>p</i>
	Controls Aβ- (n = 15) mean ± SD	Stable Aβ+ (n = 15) mean ± SD	Progressors Aβ+ (n = 15) mean ± SD	Controls Aβ- (n = 14) mean ± SD	Stable Aβ+ (n = 12) mean ± SD	Progressors Aβ+ (n = 8) mean ± SD		
FR <sub>[0-48]</sub>	29.8 ± 7.2	29.8 ± 4.8	25.5 ± 6.9	34.0 ± 7.1	31.3 ± 5.0	15.5 ± 5.7	F <sub>10,178.1</sub> = 7.3	<0.001***,a,1,2
DFR <sub>[0-48]</sub>	11.3 ± 1.8	11.2 ± 2.3	9.9 ± 2.8	11.9 ± 3.0	11.6 ± 1.6	3.0 ± 2.7	F <sub>10,180.9</sub> = 10.4	<0.001***,a,1,2
TR <sub>[0-48]</sub>	45.9 ± 2.2	46.1 ± 1.9	44.7 ± 2.5	46.6 ± 2.3	46.3 ± 2.0	34.9 ± 7.1	F <sub>10,176.4</sub> = 8.3	<0.001***,a,1,2
DTR <sub>[0-48]</sub>	15.7 ± 0.6	15.6 ± 0.6	15.4 ± 0.7	15.5 ± 0.8	15.8 ± 0.6	11.6 ± 2.8	F <sub>10,177.5</sub> = 8.8	<0.001***,a,1,2
Intrusions	1.1 ± 1.8	0.6 ± 1.0	1.6 ± 2.0	1.4 ± 3.0	1.3 ± 1.8	6.9 ± 5.6	F <sub>10,176.8</sub> = 4.4	<0.001***,a,1,2

SD, standard deviation; df, degrees of freedom; IR, immediate recall; FR, free recall; DFR, delayed free recall; TR, total recall; DTR, delayed total recall. The table represents the mean and standard deviation of each group at the first and last sessions (M<sub>0</sub>, M<sub>60</sub>, respectively) and the longitudinal interactions between group\_status\*session as a product of linear mixed effect models (lmer). Significance is reported for each outcome variable based on the following *p*-values: \*\*\**p* < 0.001; \*\**p* ≤ 0.01; \**p* ≤ 0.05.

<sup>a</sup>Cases where interactions were significant and further assessed with emmeans post-hoc analysis (corrected for multiple comparisons): <sup>1</sup>[Aβ-/controls] > [Aβ+/progressors (preclinical-AD)]: FR (*p* ≤ 0.001, M24 – M60), DFR (*p* ≤ 0.001, M24 – M60), TR (*p* ≤ 0.001, M48 – M60), DTR (*p* ≤ 0.001, M48 – M60), intrusions (*p* ≤ 0.001, M48 – M60). <sup>2</sup>[Aβ+/stable] > [Aβ+/progressors (preclinical-AD)]: FR (*p* ≤ 0.01, M24 – M60), DFR (*p* ≤ 0.05, M24 – M60), TR (*p* ≤ 0.001, M48-M60), DTR (*p* ≤ 0.001, M48 – M60), intrusions (*p* ≤ 0.001, M48 – M60).

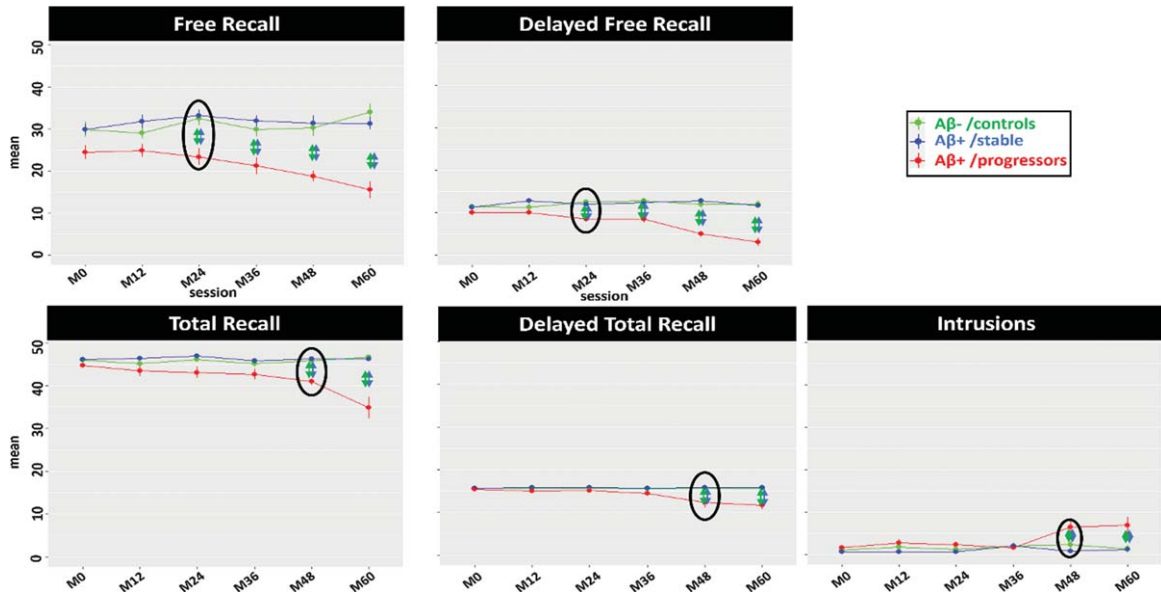


Fig. 1. Longitudinal progression of the Free and Cued Selective Reminding test [FCSRT] main scores: Between-group comparison of the mean of each main score per session (max  $\leq 48$ , excluding intrusions). Black circles represent the first session in which significant differences were noted between groups. Green arrows represent significant differences between the A $\beta$ +/progressors and the A $\beta$ -/controls; blue arrows represent significant differences between the A $\beta$ +/progressors and the A $\beta$ +/stable.

of the MBT and FCSRT to become complementary tools of AD diagnosis, with the MBT detecting emergent subtle EM alterations at an asymptomatic at-risk phase, and the FCSRT characterizing stages of EM decline from a preclinical phase.

On the FCSRT, EM decline began in the A $\beta$ +/progressors (preclinical-AD) when compared with the other two groups on both immediate and delayed recall versions, first on FR, up to 3 years prior to diagnosis (session M24), and then on TR 2 years later ( $\pm 1$  year prior to diagnosis), ex-aequo with the first production of intrusions. When in isolation, FR alterations (i.e., retrieval impairments) may result exclusively from executive or attention deficits that are not necessarily AD-linked [5–7, 10, 23, 26]. It is only when TR alterations (i.e., storage deficits) rooted in the hippocampus are also detected that EM decline can be conclusive of AD (Fig. 1) [6, 7, 10, 26]. Moreover, FCSRT-FR and TR decline (the only official AD diagnosis scores) vary highly inter-individually [5, 10]. To account for such variation and with the aim of sequentially classifying the preclinical stages of AD-linked EM decline, Grober et al. created the Stages of Objective Memory Impairment (SOMI) based solely on these scores [5, 10]. This 5-level system appears to have a remarkable potential to classify the stages of EM decline from a preclinical phase onwards and predict years to diagnosis, spanning from intact mem-

ory ( $\pm 7$  years to diagnosis) to storage impairments compatible with dementia [5, 10, 37, 38]. However, currently the SOMI is based on the FCSRT picture version (FCSRTp), known to yield higher cut-off scores than the word version (FCSRTw) used here [5, 10, 26, 38–40]. In order to obtain a more detailed characterization of preclinical progression of EM decline, we stress the need for a SOMI scale adapted to the FCSRTw version. Based on previous research and studies with large cohorts like ours, we believe this will be an asset in the current clinical context. The SOMI scale confirms the value of the FCSRT, which, being based only on FR/TR variations, might lessen the reliance on expensive biomarker testing to classify preclinical stages of AD. Nonetheless, by using only one list of words that can be easily remembered, the FCSRT prevents variations in maximum recall and TR is prone to ‘ceiling levels’ at early preclinical stages [2, 5, 6, 10, 16, 19, 21, 23]. The clinical value of the FCSRT-TR is then constrained to late preclinical stages, when the aid of semantic cues starts failing to overcome existing retrieval deficits [2, 5, 6, 10, 16, 19, 21, 23].

The MBT by using two lists of words, overcomes this limitation by introducing binding, suggested as a more sensitive measure to subtle EM decline while individuals remain cognitively normal according to FCSRT psychometric parameters [16, 19, 27–29].

Table 3  
The Memory Binding Test (MBT) main scores at baseline and after 5 years of follow-up

score_name		M <sub>0</sub>			M <sub>60</sub>			F <sub>df1,df2</sub> = value	group_status* <sup>sess</sup> p
		Controls	Stable	Progressors	Controls	Stable	Progressors		
		Aβ <sup>-</sup> (n = 15) mean ± SD	Aβ <sup>+</sup> (n = 15) mean ± SD	Aβ <sup>+</sup> (n = 15) mean ± SD	Aβ <sup>-</sup> (n = 13) mean ± SD	Aβ <sup>+</sup> (n = 12) mean ± SD	Aβ <sup>+</sup> (n = 8) mean ± SD		
Free recall	FR <sub>[0-32]</sub>	15.9 ± 4.8	18.6 ± 3.5	12.4 ± 4.2	17.8 ± 5.6	18.3 ± 6.4	8.8 ± 5.0	F <sub>10,176.1</sub> = 2.9	0.002** <sup>a,1,2</sup>
	DFR <sub>[0-32]</sub>	17.0 ± 5.4	17.7 ± 5.0	13.0 ± 7.3	21.2 ± 4.3	21.8 ± 3.7	8.4 ± 5.2	F <sub>10,162.0</sub> = 2.9	0.002** <sup>a,1,2</sup>
Cued recall	TCR <sub>A+B[0-32]</sub>	26.0 ± 3.1	28.1 ± 3.1	23.4 ± 3.7	27.6 ± 3.3	28.6 ± 2.0	18.9 ± 5.7	F <sub>10,175.0</sub> = 4.8	<0.001*** <sup>a,1,2</sup>
	PIP <sub>[0-16]</sub>	10.7 ± 2.7	12.5 ± 2.0	8.8 ± 3.3	12.2 ± 2.7	12.8 ± 1.9	6.1 ± 3.5	F <sub>10,175.4</sub> = 3.3	<0.001*** <sup>a,1,2</sup>
	TIP <sub>[0-32]</sub>	26.1 ± 3.1	27.4 ± 2.2	24.3 ± 4.1	27.9 ± 3.6	28.7 ± 2.4	19.8 ± 6.1	F <sub>10,159.6</sub> = 4.4	<0.001*** <sup>a,1,2</sup>
	DTIP <sub>[0-32]</sub>	26.2 ± 3.3	27.3 ± 2.4	24.1 ± 4.1	28.2 ± 3.0	28.8 ± 2.2	18.4 ± 5.9	F <sub>10,160.1</sub> = 5.2	<0.001*** <sup>a,1,2</sup>
Source recall	SMR <sub>[0-100%]</sub>	94.2 ± 9.7	87.0 ± 14.4	87.0 ± 16.8	95.1 ± 4.8	98.1 ± 2.7	78.6 ± 11.2	F <sub>10,161.7</sub> = 2.7	0.005** <sup>a,1,2</sup>
	DSMR <sub>[0-100%]</sub>	87.7 ± 12.7	91.7 ± 8.8	85.1 ± 14.4	93.6 ± 5.3	91.5 ± 13.4	69.8 ± 16.9	F <sub>10,163.7</sub> = 1.8	0.06 <sup>a,1,2</sup>
Intrusions	T <sub>intr</sub>	3.8 ± 4.7	1.7 ± 1.8	4.1 ± 3.2	2.2 ± 2.2	1.0 ± 0.8	9.6 ± 7.4	F <sub>10,175.0</sub> = 3.6	<0.001*** <sup>a,1,2</sup>
	Prior <sub>int</sub>	1.4 ± 1.7	0.5 ± 0.9	1.9 ± 2.1	0.4 ± 0.7	0.1 ± 0.3	1.3 ± 1.4	F <sub>10,176.2</sub> = 1.7	0.08
	Extra <sub>int</sub>	2.3 ± 3.6	1.1 v 1.5	1.9 ± 1.6	1.5 ± 1.6	0.8 ± 0.8	7.1 ± 7.0	F <sub>10,174.8</sub> = 3.8	<0.001*** <sup>a,1,2</sup>
	Ex <sub>cat_int</sub>	0.1 ± 0.3	0.1 v 0.3	0.2 ± 0.6	0.2 ± 0.6	0.1 ± 0.3	1.3 ± 1.6	F <sub>10,180.4</sub> = 1.2	0.28

SD, standard deviation; df, degrees of freedom; TCR<sub>A+B</sub>, total cue recall (paired condition); PIP, number of pairs cued recall in the paired condition; TIP, total number of items recalled in the paired condition; FR, free recall; SMR, source memory recall (%); DFR, delayed free recall; DTIP, delayed total number of items recalled in the paired condition; DSMR, delayed source memory recall (%); T<sub>int</sub>, total number of intrusions; Prior<sub>int</sub>, prior list of intrusions; Extra<sub>int</sub>, extra list of intrusions; Ex<sub>cat\_int</sub>, extra category of intrusions. The table represents the mean and standard deviation of each group at the first and last sessions (M0, M60, respectively) and the longitudinal interactions between group\_status\*session as a product of linear mixed effect models (lmer). Significance is reported for each outcome variable based on the following p-values: \*\*\*p < 0.001; \*\*p ≤ 0.01; \*p ≤ 0.05. <sup>a</sup>Cases where interactions were significant and further assessed with emmeans post-hoc analysis (corrected for multiple comparisons): <sup>1</sup>[Aβ<sup>-</sup>/controls]>[Aβ<sup>+</sup>/progressors (preclinical-AD)]: FR (p ≤ 0.01, M24 – M60), DFR (p ≤ 0.01, M24 – M60), TCR<sub>A+B</sub> (p ≤ 0.001, M24 – M60), PIP (p ≤ 0.001, M24 – M60), TIP (p ≤ 0.01, M24 – M60), DTIP (p ≤ 0.05, M12 – M60), SMR (p = 0.01, M60), DSMR (p < 0.001, M60), T<sub>int</sub> (p ≤ 0.01, M48 – M60), extra<sub>int</sub> (p ≤ 0.001, M60). <sup>2</sup>[Aβ<sup>+</sup>/stable]>[Aβ<sup>+</sup>/progressors (preclinical-AD)]: FR (p ≤ 0.01, M24 – M60), DFR (p ≤ 0.05, M12 – M60), TCR<sub>A+B</sub> (p ≤ 0.001, M24 – M60), PIP (p ≤ 0.001, M24 – M60), TIP (p ≤ 0.001, M24 – M60), DTIP (p ≤ 0.001, M24 – M60), SMR (p = 0.01, M60), DSMR (p = 0.01, M60), T<sub>int</sub> (p ≤ 0.001, M60), extra<sub>int</sub> (p ≤ 0.001, M60).

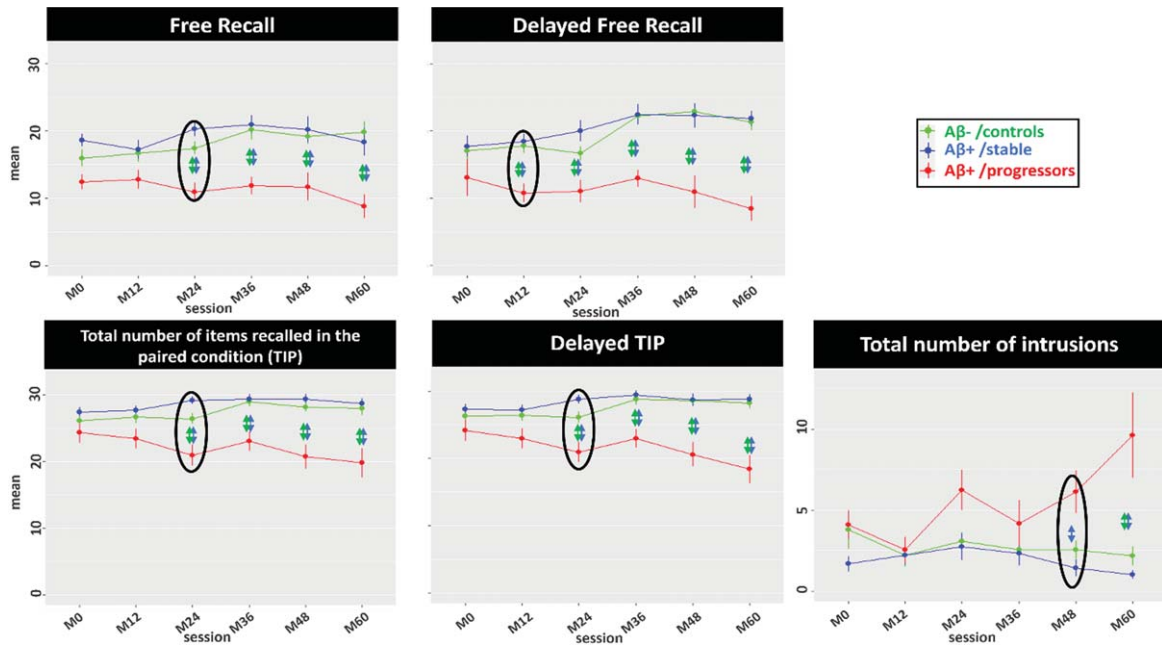


Fig. 2. **Longitudinal progression of the memory binding test [MBT] main scores:** Between-group comparison of the mean of each main score per session (max  $\leq 32$ , excluding intrusions). Black circles represent the first session in which significant differences were observed between groups. Green arrows represent significant differences between the  $A\beta+$ /progressors and the  $A\beta-$ /controls; blue arrows represent significant differences between the  $A\beta+$ /progressors and the  $A\beta+$ /stable.

Our results seem to attest this hypothesis. While in the FCSRT memory alterations appeared only up to 1 year prior to diagnosis in the  $A\beta+$ /progressors (preclinical-AD), in the MBT AD-specific alterations started 3 to 4 years prior to diagnosis in memory retention, storage, and binding (Fig. 2). Consolidation and binding processes are particularly sensitive to the appearance of subtle EM decline, underlying the typical neural correlate of AD-AHS likely translating the initial failure of the medial temporal lobe function [6, 8, 11, 16, 19, 29, 41]. Binding decrements in particular, have been shown to be highly predictive of increased progression to AD in the presence of biomarkers, and able to distinguish normal cognitive aging from amnesic mild cognitive impairment and AD dementia [11, 16, 19, 25, 28, 29]. This is likely the case of the  $A\beta+$ /progressors (preclinical-AD) respectively known to be on the path to AD.

We did not find any differences specific to  $A\beta$  burden. Only a tendency was detected up to 5 years before diagnosis for lower storage-linked scores (TCR and PIP) in the  $A\beta+$ /progressors (preclinical-AD) versus the  $A\beta+$ /stable, but not in comparison with the  $A\beta-$ /controls, as would intuitively be expected. However, this is not entirely unexpected since a considerable body of evidence suggests that

high education levels and/or an active social life in elders at-risk for AD ( $A\beta$  burden) instigates a cognitive reserve that acts as a compensatory mechanism capable of delaying evidence of the onset of a multimodal system failure [42–44]. It is possible that cognitive reserve in  $A\beta+$ /stable subjects (who already present AD biomarkers), potentially restores cognitive performance until finally a threshold is reached, and the brain loses the capacity to overcome  $A\beta$ -associated neural decline [42–44].

One advantage of this study was its vast multimodality in which the comprehensive amount of data collected permitted the creation of three distinct groups controlled for demographic (age, education level, and sex) and PET biomarkers ( $A\beta$  burden, and neurodegeneration), which in turn enabled the effect of  $A\beta$  burden and alterations linked to conversion to AD to be isolated [3]. As a longitudinal monocentric study, we were able to control variation in data collection and characterize the emergence of subtle AD-linked EM decline for 5 years, while individuals remained cognitively normal. Moreover, the use of two verbal EM tests (FCSRT and MBT) implementing controlled learning and cue recall to ensure encoding specificity and maximized recall warrants the comparison of their diagnostic capacity based

on the cross-validity of their main scores [5, 18, 19, 23–25, 27]. The distinctive advantage of the MBT, is that it increases the number and sensitivity of scores characteristic of the AD-AHS clinical phenotype, through the use of two lists of words adding ‘binding’ measures [5, 16, 18, 19, 41].

Despite its strengths, this study presents some limitations. Although covariates were controlled to minimize their influence, residual effects might have remained on the outcome measures. Moreover, the high education level of the sample is not fully representative of the general population and warrants caution when extending its conclusions. Also, we decided to retrospectively select a smaller sample of 45 individuals, but well-matched in terms of confounder factors to the 15 progressors to prodromal AD in order to, at the cost of some statistical power, reach more meaningful outcomes in this exploratory study. This can be seen as both a strength and a limitation. Nevertheless, this remains a unique opportunity to longitudinally follow such a matched sample so thoroughly before diagnosis.

In this study, the MBT displayed a considerable capacity to anticipate to an asymptomatic at-risk phase, while individuals are considered within the normal FCSRT psychometric range, the detection of otherwise imperceptible EM alterations specific to AD-AHS. The MBT increases the sensitivity of AD diagnosis by offering a higher number of AD-specific (5MBT versus 3FCSRT), and by anticipating diagnosis 1–3 years before the FCSRT. However, the FCSRT remains highly relevant as a neurocognitive tool to classify the stages of EM decline from a preclinical phase and in the prediction of years to clinical diagnosis, based solely on the classification of FR and TR scores, without requiring confirmation of the presence of AD biomarkers.

Hence, we argue that the MBT and the FCSRT are affordable neurocognitive tools, easily applicable, with good test-retest reliability, that should be complementary in AD diagnosis [2, 5, 16, 19]. The MBT as a detection tool that anticipates diagnosis of AD-linked EM deficits to an at-risk phase; the FCSRT, based on the potential added value of the SOMI scale, as a classification tool distinguishing stages of EM decline and predicting years to clinical diagnosis at a preclinical phase. The use of both tools brings enormous advantages to reach informed clinical decisions in triage for early AD treatment or initial prevention trials, imaging tests and biomarker testing. This has the potential to foster meaningful innovative treatments that delay cognitive decline and prolong

quality of life, in a disease which so far still has no cure.

## AUTHOR CONTRIBUTIONS

Filipa Raposo Pereira (Data curation; Formal analysis; Investigation; Methodology; Writing – original draft; Writing – review & editing); Nathalie George (Data curation; Methodology; Project administration; Supervision; Validation; Writing – review & editing); Gianfranco Dalla Barba (Writing – review & editing); Bruno Dubois (Conceptualization; Funding acquisition; Resources; Validation; Writing – review & editing); Valentina La Corte (Conceptualization; Data curation; Formal analysis; Funding acquisition; Methodology; Project administration; Supervision; Validation; Writing – review & editing).

## ACKNOWLEDGMENTS

The authors would like to thank Maximilien Chaumon, Nicolas Villain, and Stephanie Bombois for their valuable perspectives on the research presented in this article.

## FUNDING

The research leading to these results received funding from the Fondation pour la Recherche sur Alzheimer (FRA, project: ePARAD) and from the program “Investissements d’avenir” ANR-10-IAIHU-06.

The INSIGHT study (IDRCB: 2012-A01731-42) was supported by INSERM (C12-59) in collaboration with Institut du Cerveau et de la Moelle Epinière (ICM), the Institut Hospitalo-Universitaire at ICM, and with funding from Pfizer and the Investissement d’Avenir (ANR-10-AIHU-06) that was used for the recruitment of clinical research assistants, neuropsychologists, and a study physician. The study was done in collaboration with the Centre Hospitalier Universitaire de Bordeaux (study number CIC EC7). The Plan-Alzheimer Foundation provided funding, and along with the MEMENTO study contributed to the costs of MRI and  $^{18}\text{F}$ -fluorodeoxyglucose PET. Avid provided the  $^{18}\text{F}$ -florbetapir ligand for amyloid- $\beta$  PET.

## CONFLICT OF INTEREST

The authors declare no other conflicts of interest.

## DATA AVAILABILITY

The data supporting the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: <https://dx.doi.org/10.3233/JAD-230921>.

## REFERENCES

- [1] Gauthier S, Webster C, Servaes S, Morais JA, Rosa-Neto P (2022) *World Alzheimer Report 2022. Life after diagnosis: Navigating treatment, care and support*. Alzheimer's Disease International, London.
- [2] Loewenstein DA, Curiel RE, Duara R, Buschke H (2018) Novel cognitive paradigms for the detection of memory impairment in preclinical Alzheimer's disease. *Assessment* **25**, 348-359.
- [3] Dubois B, Epelbaum S, Nyasse F, Bakardjian H, Gagliardi G, Uspenskaya O, Houot M, Lista S, Cacciamani F, Potier M-C (2018) Cognitive and neuroimaging features and brain  $\beta$ -amyloidosis in individuals at risk of Alzheimer's disease (INSIGHT-preAD): A longitudinal observational study. *Lancet Neurol* **17**, 335-346.
- [4] Jack CR, Wiste HJ, Weigand SD, Therneau TM, Knopman DS, Lowe V, Vemuri P, Mielke MM, Roberts RO, Machulda MM, Senjem ML, Gunter JL, Rocca WA, Petersen RC (2017) Age-specific and sex-specific prevalence of cerebral  $\beta$ -amyloidosis, tauopathy, and neurodegeneration in cognitively unimpaired individuals aged 50-95 years: A cross-sectional study. *Lancet Neurol* **16**, 435-444.
- [5] Grober E, Veroff AE, Lipton RB (2018) Temporal unfolding of declining episodic memory on the Free and Cued Selective Reminding Test in the prodementia phase of Alzheimer's disease: Implications for clinical trials. *Alzheimers Dement (Amst)* **10**, 161-171.
- [6] Sarazin M, Berr C, De Rotrou J, Fabrigoule C, Pasquier F, Legrain S, Michel B, Puel M, Volteau M, Touchon J, Verny M, Dubois B (2007) Amnesic syndrome of the medial temporal type identifies prodromal AD A longitudinal study. *Neurology* **69**, 1859-1867.
- [7] Sarazin M, Chauviré V, Gerardin E, Colliot O, Kingnéhun S, De Souza LC, Hugonot-Diener L, Garnero L, Lehericy S, Chupin M, Dubois B (2010) The amnesic syndrome of hippocampal type in Alzheimer's disease: An MRI study. *J Alzheimers Dis* **22**, 285-294.
- [8] Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, Bakardjian H, Benali H, Bertram L, Blennow K, Broich K, Cavado E, Crutch S, Dartigues JF, Duyckaerts C, Epelbaum S, Frisoni GB, Gauthier S, Genton R, Gouw AA, Habert MO, Holtzman DM, Kivipelto M, Lista S, Molinuevo JL, O'Bryant SE, Rabinovici GD, Rowe C, Salloway S, Schneider LS, Sperling R, Teichmann M, Carrillo MC, Cummings J, Jack CR (2016) Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimers Dement* **12**, 292-323.
- [9] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster M V., Phelps CH (2011) Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* **7**, 280-292.
- [10] Grober E, Qi Q, Kuo L, Hassenstab J, Perrin RJ, Lipton RB (2021) Stages of objective memory impairment predict Alzheimer's disease neuropathology: Comparison with the Clinical Dementia Rating Scale-Sum of Boxes. *J Alzheimers Dis* **80**, 185-195.
- [11] Spaan PEJ (2016) Episodic and semantic memory impairments in (very) early Alzheimer's disease: The diagnostic accuracy of paired-associate learning formats. *Cogent Psychol* **2**, 1109782.
- [12] Chételat G, La Joie R, Villain N, Perrotin A, De La Sayette V, Eustache F, Vandenberghe R (2013) Amyloid imaging in cognitively normal individuals, at-risk populations and preclinical Alzheimer's disease. *Neuroimage Clin* **2**, 356-365.
- [13] Chipi E, Salvadori N, Farotti L, Parnetti L (2019) Biomarker-based signature of Alzheimer's disease in pre-MCI individuals. *Brain Sci* **9**, 213.
- [14] Rentz DM, Parra Rodriguez MA, Amariglio R, Stern Y, Sperling R, Ferris S (2013) Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: A selective review. *Alzheimers Res Ther* **5**, 58.
- [15] Dubois B, Feldman HH, Jacova C, rey Cummings JL, DeKosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, Meguro K, Pasquier F, Robert P, Rossor M, Salloway S, Sarazin M, de Souza LC, Stern Y, Visser PJ, Scheltens P (2010) Revising the definition of Alzheimer's disease: A new lexicon. *Lancet Neurol* **9**, 1118-1127.
- [16] Buschke H, Mowrey WB, Ramratan WS, Zimmerman ME, Loewenstein DA, Katz MJ, Lipton RB (2017) Memory binding test distinguishes amnesic mild cognitive impairment and dementia from cognitively normal elderly. *Arch Clin Neuropsychol* **32**, 29-39.
- [17] Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, Dekosky ST, Gauthier S, Selkoe D, Bateman R, Cappa S, Crutch S, Engelborghs S, Frisoni GB, Fox NC, Galasko D, Habert MO, Jicha GA, Nordberg A, Pasquier F, Rabinovici G, Robert P, Rowe C, Salloway S, Sarazin M, Epelbaum S, de Souza LC, Vellas B, Visser PJ, Schneider L, Stern Y, Scheltens P, Cummings JL (2014) Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. *Lancet Neurol* **13**, 614-629.
- [18] Grober E, Lipton RB, Hall C, Crystal H (2000) Memory impairment on free and cued selective reminding predicts dementia. *Neurology* **54**, 827-832.
- [19] Gramunt N, Sánchez-Benavides G, Buschke H, Lipton RB, Masramon X, Gispert JD, Peña-Casanova J, Fauria K, Molinuevo JL (2016) Psychometric properties of the Memory Binding Test: Test-retest reliability and convergent validity. *J Alzheimers Dis* **50**, 999-1010.
- [20] Salmon D (2012) Loss of semantic knowledge in mild cognitive impairment. *Am J Psychiatry* **169**, 1226-1229.
- [21] Papp KV, Rentz DM, Orlovsky I, Sperling RA, Mormino EC (2017) Optimizing the preclinical Alzheimer's cog-

- nitive composite with semantic processing: The PACC5. *Alzheimers Dement (N Y)* **3**, 668-677.
- [22] Tounsi H, B Deweer B, Ergis AM, Van der Linden M, Pillon B, Michon A, Dubois B (1999) Sensitivity to semantic cuing: An index of episodic memory dysfunction in early Alzheimer disease. *Cogent Psychol* **13**, 38-46.
- [23] Auriacombe S, Helmer C, Amieva H, Berr C, Dubois B, Dartigues JF (2010) Validity of the free and cued selective reminding test in predicting dementia: The 3C study. *Neurology* **74**, 1760-1767.
- [24] Grober E, Buschke H, Korey SR (1987) Genuine memory deficits in dementia. *Dev Neuropsychol* **3**, 13-36.
- [25] Papp KV, Amariglio RE, Mormino EC, Hedden T, Dekhytar M, Johnson KA, Sperling RA, Rentz DM (2015) Free and cued memory in relation to biomarker-defined abnormalities in clinically normal older adults and those at risk for Alzheimer's disease. *Neuropsychologia* **73**, 169-175.
- [26] Papp K V, Rentz DM, Mormino EC, Schultz AP, Amariglio RE, Quiroz Y, Johnson KA, Sperling RA (2017) Cued memory decline in biomarker-defined preclinical Alzheimer disease. *Neurology* **88**, 1431-1438.
- [27] Buschke H (2014) *Rationale of the Memory Binding Test*. In *Dementia and memory*, Nilsson L-G, Ohta N, eds. Psychology Press, pp. 55-71.
- [28] Mowrey WB, Lipton RB, Katz MJ, Ramratan WS, Loewenstein DA, Zimmerman ME, Buschke H (2018) Memory Binding Test predicts incident dementia: Results from the Einstein Aging Study. *J Alzheimers Dis* **62**, 293-304.
- [29] Mowrey WB, Lipton RB, Katz MJ, Ramratan WS, Loewenstein DA, Zimmerman ME, Buschke H (2016) Memory Binding Test predicts incident amnesic mild cognitive impairment. *J Alzheimers Dis* **53**, 1585-1595.
- [30] Gagliardi G, Epelbaum S, Houot M, Bakardjian H, Boukadida L, Revillon M, Dubois B, Dalla Barba G, La Corte V (2019) Which episodic memory performance is associated with Alzheimer's disease biomarkers in elderly cognitive complainers? Evidence from a longitudinal observational study with four episodic memory tests (Insight-PreAD). *J Alzheimers Dis* **70**, 811-824.
- [31] Folstein MF, Folstein SE, Mchugh PR (1975) "Mini-Mental State" A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.
- [32] Morris JC (1993) The clinical dementia rating (CDR): Current version and scoring rules. *Neurology* **43**, 2412-2414.
- [33] Habert MO, Bertin H, Labit M, Diallo M, Marie S, Martineau K, Kas A, Causse-Lemercier V, Bakardjian H, Epelbaum S, Chételat G, Houot M, Hampel H, Dubois B, Mangin JF, Audrain C, Bakardjian H, Benali H, Bertin H, Boukadida L, Cacciamani F, Causse-Lemercier V, Cavado E, Chiesa P, Colliot O, Dos Santos A, Dubois B, Durrleman S, Epelbaum S, Gagliardi G, Genthon R, Habert MO, Hampel H, Jungalee N, Kas A, Lehericy S, Lamari F, Letondor C, Levy M, Lista S, Mochel F, Nyasse F, Poisson C, Potier MC, Revillon M, Rojkova K, Roy P, Santos-Andrade K, Santos A, Simon V, Sole M, Tandetnik C, Thiebaud De Schotten M (2017) Evaluation of amyloid status in a cohort of elderly individuals with memory complaints: Validation of the method of quantification and determination of positivity thresholds. *Ann Nucl Med* **32**, 75-86.
- [34] Dubois B, Slachevsky A, Litvan I, Pillon B (2000) The FAB: A frontal assessment battery at bedside. *Neurology* **55**, 1621-1626.
- [35] McNair DM, Kahn RJ (1983) Self-assessment of cognitive deficits. In *Assessment in geriatric psychopharmacology*, Crook T, Ferris S, Bartus R, eds. Mark Powley, New Canaan, CT, pp. 119-136.
- [36] Van Der Linden M, Calicis F, Adam S, and members of the GREMEM (2004) L'épreuve de rappel libre/rappel indice à 16 items (RL/RI-16). Solar Editeur.
- [37] Grober E, Wang C, Kitner-Triolo M, Lipton RB, Kawas C, Resnick SM (2021) Prognostic value of learning and retention measures from the Free and Cued Selective Reminding Test to identify incident mild cognitive impairment. *J Int Neuropsychol Soc* **28**, 292-299.
- [38] Grober E, Lipton RB, Sperling RA, Papp KV, Johnson KA, Rentz DM, Veroff AE, Aisen PS, Ezzati A (2022) Associations of stages of objective memory impairment with amyloid PET and structural MRI: The A4 Study. *Neurology* **98**, E1327-E1336.
- [39] Zimmerman ME, Katz MJ, Wang C, Burns LC, Berman RM, Derby CA, L'Italien G, Budd D, Lipton RB (2015) Comparison of "Word" vs. "Picture" version of the Free and Cued Selective Reminding Test (FCSRT) in older adults. *Alzheimers Dement (Amst)* **1**, 94-100.
- [40] Delgado C, Muñoz-Neira C, Soto A, Martínez M, Henríquez F, Flores P, Slachevsky A (2016) Comparison of the psychometric properties of the "word" and "picture" Versions of the Free and Cued Selective Reminding Test in a Spanish-speaking cohort of patients with mild Alzheimer's disease and cognitively healthy controls. *Arch Clin Neuropsychol* **31**, 165-175.
- [41] Dubois B, Albert M (2004) Amnesic MCI or prodromal Alzheimer's disease? *Lancet Neurol* **3**, 246-248.
- [42] Stern Y, Arenaza-Urquijo EM, Bartrés-Faz D, Belleville S, Cantillon M, Chételat G, Ewers M, Franzmeier N, Kempermann G, Kremen WS, Okonkwo O, Scarmeas N, Soldan A, Udeh-Momoh C, Valenzuela M, Vemuri P, Vuoksimaa E, Urquijo EMA, Cantillon M, Clouston SAP, Estanga A, Gold B, Habeck C, Jones R, Kochhann R, Lim YY, Martínez-Lage P, Morbelli S, Okonkwo O, Ossenkoppele R, Pettigrew C, Rosen AC, Song X, Van Loenhoud AC (2020) Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimers Dement* **16**, 1305-1311.
- [43] Arenaza-Urquijo EM, Vemuri P (2018) Resistance vs resilience to Alzheimer disease. *Neurology* **90**, 695-703.
- [44] Arenaza-Urquijo EM, Bejanin A, Gonneaud J, Wirth M, La Joie R, Mutlu J, Gaubert M, Landeau B, de la Sayette V, Eustache F, Chételat G (2017) Association between educational attainment and amyloid deposition across the spectrum from normal cognition to dementia: Neuroimaging evidence for protection and compensation. *Neurobiol Aging* **59**, 72-79.