Early Detection of Learning Difficulties when Confronted with Novel Information in Preclinical Alzheimer's Disease Stage 1

Adrià Tort-Merino[a], Natalia Valech[a], Claudia Peñaloza[b], Petra Grönholm-Nyman[c], María León[a], Jaume Olives[a], Ainara Estanga[d], Mirian Ecay-Torres[d], Juan Fortea[e], Pablo Martínez-Lage[d], José L. Molinuevo[a,f], Matti Laine[c], Antoni Rodríguez-Fornells[b,g,h,1] and Lorena Rami[a,i,1,*]

[a]*Alzheimer's Disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clinic, Barcelona, Spain*
[b]*Cognition and Brain Plasticity Group, Bellvitge Biomedical Research Institute- IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain*
[c]*Department of Psychology, Åbo Akademi University, Turku, Finland*
[d]*Neurología, Fundación CITA-Alzhéimer Fundazioa, Centro de Investigación y Terapias Avanzadas, San Sebastián, Guipúzcoa, España*
[e]*Memory Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau and Institute of Biomedical Research, Barcelona, Spain*
[f]*Clinical Research Program, Barcelonaβeta Brain Research Center, Pasqual Maragall Foundation, Barcelona, Spain*
[g]*Department of Cognition, Development and Education Psychology, Campus Bellvitge, University of Barcelona, L'Hospitalet de Llobregat, Barcelona, Spain*
[h]*Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain*
[i]*August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona, Spain*

Abstract. We employed a highly demanding experimental associative learning test (the AFE-T) to explore memory functioning in Preclinical Alzheimer's Disease stage 1 (PreAD-1) and stage 2 (PreAD-2). The task consisted in the learning of unknown object/name pairs and our comprehensive setup allowed the analysis of learning curves, immediate recall, long-term forgetting rates at one week, three months, and six months, and relearning curves. Forty-nine cognitively healthy subjects were included and classified according to the presence or absence of abnormal CSF biomarkers (Control, $n = 31$; PreAD-1, $n = 14$; PreAD-2, $n = 4$). Control and PreAD-1 performances on the experimental test were compared by controlling for age and education. These analyses showed clear learning difficulties in PreAD-1 subjects (F = 6.98; $p = 0.01$). Between-group differences in long-term forgetting rates were less notable, reaching statistical significance only for the three-month cued forgetting rate (F = 4.83; $p = 0.03$). Similarly, relearning sessions showed only statistical trends between the groups (F = 3.22; $p = 0.08$). In the whole sample, significant correlations between CSF $A\beta_{42}$/tau ratio and the AFE-T were found, both in the total learning score ($r = 0.52$; $p < 0.001$) and in the three-month cued forgetting rate ($r = -0.38$; $p < 0.01$). Descriptive subanalyses involving PreAD-2 suggested greater learning and recall difficulties in these subjects when compared with the PreAD-1 group. The present results suggest that explicit learning difficulties when binding information could be one of the earliest signs of the future emergence of episodic memory difficulties on the Alzheimer's disease continuum. Our findings indicate that the AFE-T is a sensitive test, capable of detecting subtle memory difficulties in PreAD-1.

Keywords: Alzheimer's disease, biomarkers, cognitive aging, memory, neuropsychology

---

[1]These authors contributed equally to this work.

*Correspondence to: Lorena Rami, PhD, Alzheimer's Disease and Other Cognitive Disorders Unit, Neurology Service, Hospital Clínic, IDIBAPS, Villarroel 170, 08036 Barcelona, Spain. Tel.: +34 932275785; Fax: +34 932275783; E-mail: lrami@clinic.cat.

INTRODUCTION

The identification of subtle cognitive changes in preclinical Alzheimer's disease (Pre-AD) has long been considered critical for predicting progression toward later clinical AD stages. Within the Pre-AD phase, three preclinical stages were defined by the National Institute of Aging and Alzheimer's Association (NIA-AA): stage 1 with abnormal amyloid-β (Aβ) levels, stage 2 with both amyloidosis and neurodegeneration (including elevated levels of CSF tau or brain atrophy), and stage 3 with the onset of subtle cognitive decline [1, 2]. The usefulness of the NIA-AA staging has been demonstrated in recent reports involving Pre-AD subjects [3–6]. These studies on the different Pre-AD stages support the idea that the co-occurrence of Aβ deposition and neurodegeneration (i.e., Pre-AD-2) accelerates cognitive decline in cognitively healthy individuals and is needed for the emergence of subtle cognitive difficulties. Another important measure in Pre-AD studies is the apolipoprotein E (APOE) ε4 genotype. In that line, Lim et al. [7] analyzed the cognitive performance of 144 healthy older adults classified as APOE ε4 carriers ($n = 61$) and APOE ε4 noncarriers ($n = 83$). They found a moderate negative relationship between cerebral Aβ and episodic memory performance only in APOE ε4 carriers.

A cognitive feature of Pre-AD subjects is that all of them, regardless of staging, have scores within the normal range on standard neuropsychological tests. Thus, most studies using a cross-sectional design have failed to find a relationship between cognitive performance on standard neuropsychological tests and biomarker evidence of AD in clinically asymptomatic at-risk individuals [8–11]. However, studies such as the conducted by Rentz et al. [12] found associations between a high demanding face-name associative memory test and Aβ accumulation in brain regions associated with memory systems. Other later studies including Pre-AD staging, reported group differences between PreAD-2 subjects and controls but failed to find cross-sectional difficulties in PreAD-1 [4, 5], and only a posterior report showed that a demanding memory test (the free recall subtest of the Memory Capacity Test; MCT) managed to discriminate between PreAD-1 subjects and controls [6]. The MCT is a high demanding associative memory task consisting on binding a total of 32 words (distributed in two lists) with a semantic cue to improve encoding and recall. The test includes free and cued immediate recall and free and cued delayed recall at 30 minutes. Therefore, it seems mandatory to develop more sensitive cognitive measures to detect subtle cognitive difficulties at Pre-AD stages, especially in PreAD-1.

In the present study, we adapted an innovative associative learning task based on the Ancient Farming Equipment (AFE) paradigm [13] to assess learning, recall, and relearning in Pre-AD subjects. This task, originally devised to examine the early stages of learning new words in one's native tongue, engages the declarative memory system in order to properly associate unfamiliar names to objects that are equally novel. The fact that participants need to create a new associative link between the representations at the lexical (new-word trace) and at the visual-conceptual level (new-object) makes this task highly demanding, especially compared to classical episodic memory tasks that require solely the memorization of words or existing objects. Task difficulty is further enhanced by the fact that AFE performance is measured by spontaneous oral production of the novel word. Based on the influential Complementary Learning Systems model (CLS) [14], it has been hypothesized that the initial encoding of a new word and its associative link to a new picture (concept) engages medial-temporal lobe regions (e.g., hippocampal and parahippocampal cortices) [13, 15–18]. There is evidence showing that these regions are affected early in AD [19–21].

To date the AFE paradigm has been mainly used to study acquisition of new words in healthy individuals [22–24] and in two studies concerning mild cognitive impairment (MCI) and AD [25, 26]. A recent fMRI study using the AFE paradigm in an aphasic patient and in healthy controls showed a clear involvement of medial temporal lobe regions during the learning period [21]. In a later study, Grönholm et al. [25] studied the neural correlates of the AFE paradigm in MCI subjects using positron emission tomography (PET). Compared to age-matched controls, MCI subjects showed increased activation in the anterior cingulate cortex, suggesting that the naming of newly learned objects imposed additional executive and attentional demands. The behavioral results of this study showed learning differences between controls and MCI since the first training run, indicating that initial learning measures were sensitive to MCI. The same authors reported learning and forgetting differences between MCI, AD, and controls in a later study using the same paradigm [26]. These results showed that both learning and forgetting performances were significantly impaired in the MCI subjects in comparison

to controls but all groups showed similar forgetting patterns, and that the MCI group benefited less from phonological cueing than controls.

Because of the novelty of the learning materials in the test (here coined as the AFE-Test, or AFE-T) and the high demands of its outcome measure (spontaneous naming), participants need multiple runs in order to be able to learn the set of new object/word pairs. It has been suggested that learning across multiple trials may provide the most sensitive index for initial diagnosis of MCI [27]. Besides, a crucial difference between this test and other standard memory tests is the longer time span for assessing forgetting (which comprised three time-points: one week, 3 months, and 6 months after initial learning). Previous memory tasks used in MCI and in early detection of people at risk of AD have usually evaluated delayed memory recall or recognition only after a 20-30-min delay from the encoding phase [28]. Importantly for the present research, recent studies have suggested that longer-term follow-up is crucial for tracking forgetting rates, including both recognition and recall, in order to obtain a level of sensitivity able to detect subtle memory difficulties [29, 30]. In addition, we included a phonological cueing test in the long-term follow-up in order to evaluate recall processes in more detail. Finally, the AFE-T also includes an additional *relearning task* carried out six months after the initial learning. Some studies suggest that information that has become inaccessible in recall or recognition tests can be reactivated by relearning tasks [31–33]. In sum, the information provided by the present comprehensive AFE-T includes detailed learning curves (learning rate), short and long-term forgetting measures, and a relearning curve.

We expected that the AFE-T, being a highly demanding learning and memory test, would enable us to detect subtle difficulties in learning and/or recall in Pre-AD subjects that cannot be detected by standard neuropsychological tests. Moreover, we aimed to evaluate the possible relationship between learning and memory performance and CSF proteins in Pre-AD subjects.

MATERIAL AND METHODS

Participants

Forty-nine cognitively normal subjects were included in the present study between 2013 and 2015. They were recruited from longitudinal ongoing projects at three Spanish memory centers: Hospi-

tal Clinic (HC), Hospital de la Santa Creu i Sant Pau (HSP) in Barcelona, and the CITA-Alzheimer Foundation (CITA) in San Sebastian. All subjects were bilingual (Catalan-Spanish for HC and HSP, and Basque-Spanish for CITA participants). The ethics committee of the Hospital Clinic of Barcelona approved the study, and all participants provided signed, informed consent before undergoing the neuropsychological assessment, MRI, and the lumbar puncture. All subjects had to meet the following inclusion criteria: a) at least three years of formal education, so as to exclude mental retardation or congenital learning disability, b) Mini-Mental State Examination (MMSE) [34] score >24, and c) objective cognitive performance within the normal range (cutoff 1.5 SD from normative mean) in all tests on a specific neuropsychological battery (see below). The following exclusion criteria were applied: a) presence of any neurological diagnosis, b) presence of a serious or unstable medical condition, c) diagnosis of a major psychiatric disorder including schizophrenia, major depression or substance abuse, and d) presence of a CSF pattern compatible with suspected non-amyloid pathology (SNAP). In accordance with the guidelines proposed by the National Institute on Aging and the Alzheimer's Association (NIA-AA) for defining Pre-AD for research purposes [1], healthy subjects were classified into three groups: control (CTR, $n = 31$), preclinical Alzheimer disease stage 1 (PreAD-1, $n = 14$) and preclinical Alzheimer's disease stage 2 (PreAD-2, $n = 4$). CSF tau and p-tau levels and MRI imaging (evaluated by an expert neurologist in order to exclude cases with brain structural damage or hippocampal atrophy) were used to classify Pre-AD subjects into PreAD-1 or PreAD-2.

Determination of biological and CSF biomarkers

All subjects underwent a lumbar puncture between 9 a.m. and 12 p.m. In the extraction, 10 ml of CSF was collected. The samples were centrifuged and stored in polypropylene tubes at –80°C within the first hour after extraction. CSF $A\beta_{42}$ levels, total tau (tau), and phosphorylated tau at threonine-181 (ptau) were measured by enzyme-linked immunosorbent assay kits (Innogenetics, Ghent, Belgium). Cut-off values of abnormality for each CSF biomarker were defined according to previous work [35, 36]: a) $A\beta_{42} \leq 550$ pg/ml, b) tau ≥ 350 pg/ml for subjects younger than 50 years, ≥ 400 pg/ml for subjects between 50–70 years old, and ≥ 450 pg/ml for subjects older than 70 years, and c) ptau ≥ 75 pg/ml. The time lapse between

the lumbar puncture and the AFE-T assessment was 2.25 (1.4) years. Both, the AFE-T administrator and the study participants were blind to the CSF results.

Apolipoprotein E analysis

Genomic DNA was extracted from peripheral blood of probands using the QIAamp DNAblood minikit (Qiagen AG, Basel, Switzerland). ApoE genotyping was performed by polymerase chain reaction amplification and HhaI restriction enzyme digestion. The study participants were blind to the ApoE results.

Neuropsychological battery and psychological assessment

All participants were assessed with a comprehensive neuropsychological battery, administered by a trained neuropsychologist blind to the CSF results. The battery encompassed four cognitive domains. The memory domain included the Free and Cued Selective Reminding Test [37], the language domain comprised of the Boston Naming Test [38] and Semantic fluency [39]; the visual perception domain contained the number location subtest of the VOSP battery [40], and the executive functions domain consisted of the Trail Making Test [41], the Stroop Test [42], the Symbol Digit Modalities Test [43], and the Digit Span test of the WAIS [44]. Global cognition was assessed with the MMSE [34]. Premorbid intelligence was assessed with the Spanish word accentuation test [45]. The average time lapse between the neuropsychological assessment and the AFE-T administration was 1.21 (0.2) months.

Word and pseudoword spans

At the end of the one-week recall session of the AFE-T, two word and pseudoword verbal span tests were administered. We developed these experimental tests to assess verbal working memory in the context of both familiar and unknown words, and to ensure that learning performance between the groups in the AFE-T was not influenced by different working memory capacities. In both tests, words or pseudowords appeared one at a time for 3 s on a white background on a computer screen. The participants were asked to read each word aloud and try to remember them in the exact order. After the items were presented, an image of a microphone appeared on the screen and the participants were asked to repeat all the items just presented in exactly the same order. Points were given for fully correct responses (i.e., when the exact words were pronounced in the exact order of presentation). When a correct response was given on at least one out of the three sequences of a given span length, the next series with a higher length was presented. The task was initiated with two-item sequences and ended with a maximum of eight items for the word span and six for the pseudoword span. The task was interrupted if the participant was unable to repeat any of the three sequences of a given span. The total score corresponded to the maximum span that the participant was able to repeat correctly.

The ancient farming equipment test (AFE-T)

The task was to learn to orally name new object/name pairs. The objects were 24 black-and-white images of ancient farming equipment taken from the AFE paradigm [13]. These objects are unknown today (see an example of a novel picture in Fig. 1), and subjects' unfamiliarity with the object was confirmed in a pre-training screening test. In this screening test, the objects were presented one by one and the participant was requested to indicate whether they knew them. Each object was paired with a pseudoword, that is, a non-existing word that follows the phonotactic rules of Spanish [46]. The object names consisted of 14 bisyllabic and 10 trisyllabic pseudowords that did not exist in the Spanish dictionary (e.g., *gorsi, folute*; see the complete list in Appendix 1). All the stimuli were presented on a computer screen against a white background using the E-prime 2.0 (Psychology Software Tools, Inc., PA,
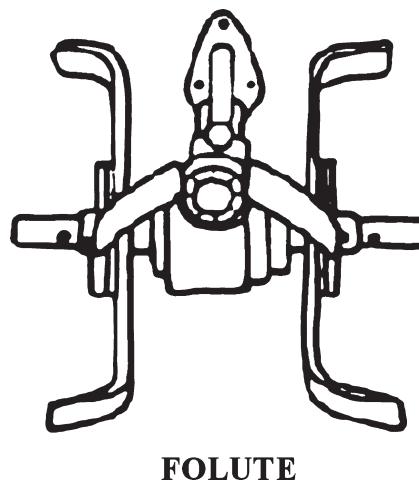

FOLUTE

Fig. 1. Example of a novel picture and a novel word in the AFE-T (see Appendix 1 for a full list of materials used).
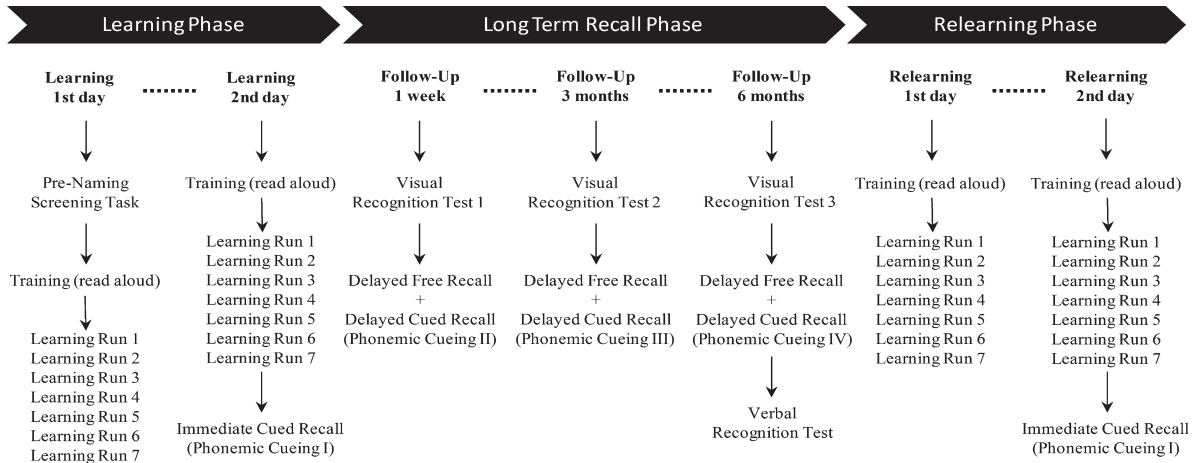
Fig. 2. Schematic design of the phases of the AFE-T.

USA). In order to thoroughly explore learning and forgetting in Pre-AD, the test design included two consecutive learning sessions, one immediate cued recall, three long-term delayed recall/recognition sessions, and two relearning sessions. The AFE-T had a total duration of six months. All the phases are explained in detail below (see Fig. 2 for a schematic description of the overall design used).

Initial learning sessions (LS)

Two learning sessions were performed on two consecutive days (LS1 for the first day, and LS2 for the second day). Each learning session included a total of seven runs and took approximately 45 min. Before starting the learning phase, each of the 24 object/name pairs was displayed for 7 s with a 500 ms pause between them. The participants were asked to read aloud the name of the object printed below, and to try to learn each object/name pair. After the presentation, the seven learning runs were performed. In each run, the participants were presented with the objects one at a time, and were asked to spontaneously say its name aloud. They were given a maximum of 7 s to recall the name of each object. After this, the correct name of the object appeared below the object for 4 s, regardless of whether the participant was able to say the correct name. The following object was presented after a 500 ms interval. The order of presentation of the objects in each run was randomized.

Immediate Cued Recall (ICR)

After the last run of the LS2, an immediate cued-recall test was administered. In this task, each

object was presented one at a time. When the object appeared, the experimenter verbally provided the first syllable of the object's name (phonemic cue). The participant then had a maximum of 7 s to provide the complete correct name. In this run, the administrator did not give the whole correct name.

Long-term recall

Long-term recall was examined at one week, three months, and six months after the initial learning phase. Each session took 10–15 min and began with a visual recognition task. The visual recognition task required the participant to identify the 24 trained objects among 24 foils (maximum score of 48). The stimuli were presented one by one for 7 s, with a 500 ms interstimulus interval, in a pseudo-randomized order. The participant had to verbally respond "YES" or "NO" to indicate whether the object had been among the 24 trained items. The recognition task was followed by free recall. Here, each trained object appeared on the screen in a randomized order, and the participant was asked to name it orally. A maximum of 7 s was given to name each object. When the participant could not provide the correct response, the experimenter provided the first syllable of the name (*delayed cued recall*). The same procedure of cued recall was repeated one week, three months, and six months after the initial learning. At the end of the 6-month period, a picture-word matching task was administered in order to further explore the participants' word acquisition through recognition memory. In this task, three pseudowords, the target and two foils sharing the

same initial syllable, were presented beside each object. The participants were requested to choose the correct alternative for each stimulus (maximum score of 24).

Relearning phase (RL)

Relearning started immediately after the end of the 6-month session, following the verbal recognition memory task. Relearning was also performed on two consecutive days, and followed exactly the same procedure as in the learning phase.

Scoring system

All verbal responses were recorded for offline scoring. Following the scoring procedure of the AFE paradigm, a response was considered correct (score = 1) when: (a) the participant recalled the exact name of the object, or when (b) the name recalled differed by a single phoneme from the original name. Under (b), the following cases were considered: the substitution, addition, or omission of a single phoneme at any given position of the word, or a change in position of an otherwise correct phoneme. This criterion was applied for all runs. Thus, for each run, the range for the scores was 0–24.

Statistical analyses

Statistical analyses were performed using the SPSS (v.22.0) package for Windows. In all analyses, a $p < 0.05$ was considered to be significant. The main analyses were performed comparing the CTR and the PreAD-1 groups. Demographical data, levels of CSF $A\beta_{42}$, CSF tau, and CSF ptau, and APOE $\varepsilon4$ frequencies were compared using Student t-tests for independent samples and χ^2 analyses when appropriate.

AFE-T learning and relearning scores were analyzed using mixed-model analyses of variance, controlling for age and years of education. In these analyses, the within-group learning curves (runs), overall group differences (group), and the interaction between learning and group (run × group) were explored. Additionally, analyses of covariance (ANCOVA) controlling for age and years of education with *post-hoc* Bonferroni corrections were performed to analyze the specific runs in which the scores differed significantly between the two groups.

Forgetting rates for CTR and PreAD-1 groups were determined to analyze the delayed recall scores relative to the learning scores of each group. The forgetting rate was defined as one minus the ratio between each delayed session score and the score obtained on the last learning run (e.g., 1-(one-week free recall score/LS2 run 7 score), for one-week free forgetting rate). In this way, the forgetting rate represents the mean percentage of object/name items previously learned that were forgotten. In the delayed recall sessions, the free and cued raw scores, the forgetting rates and the recognition scores were compared between the groups using ANCOVAs adjusted for age and education. Finally, to explore the possible relearning benefits (i.e., the reactivation of stored information that cannot be voluntarily recalled), paired *t*-tests were run comparing the within-group scores of the first runs of the learning versus relearning sessions.

Due to the fact that the present AFE-T version has not been used before, our study is essentially explorative. Following the recommendations by Armstrong [47], we applied Bonferroni corrections because here "a large number of tests are carried out without pre-planned hypothesis in an attempt to establish any results that may be significant" (op.cit., p. 505).

Using the whole sample, Pearson bivariate correlations were calculated to assess the associations between the CSF $A\beta_{42}$ levels and the following AFE-T scores: total learning and three-month cued forgetting rate. Total learning, operationalized as the total score in the last run of the learning sessions, was hypothesized to reflect mainly acquisition. It has been suggested that the latter trials in repetitive tests of memory are more strongly related to the integrity of medial temporal lobe structures, whereas early trials correlate more strongly with inferior parietal, middle frontal gyrus, and temporal pole regions of interest [48]. Three-month cued forgetting rate was hypothesized to represent consolidation in long-term memory. The cued nature of recall allows for a more specific measure of consolidation, minimizing the role of executive components in recall [49]. Cued recall at three months was hypothesized to be the best measure to capture the longer term forgetting rate while avoiding floor effect.

Finally, even though the subjects had to be within the normal range on all the neuropsychological tests administered to be eligible, we ran t tests to explore possible between-group differences on these tests. For these analyses, the scaled scores (i.e., Neuronorma) of each test were used.

Table 1
Demographics, biological data, and CSF biomarker levels of CTR and PreAD-1 groups

Parameters	CTR ($n=31$)	PreAD-1 ($n=14$)	t	p value
Demographics				
Gender (% women)	70.9%	78.6%	0.29[a]	0.59
Age	64.8 ± 6.4 [49–77]	67.8 ± 7.1 [58– 78]	1.39	0.17
Years of education	11.6 ± 3.7	10.8 ± 3.9	–0.69	0.49
Biological data and CSF				
APOE Ɛ4 (% positive)	6.5%	57.1%	14.34[a]	<0.0001**
Aβ_{42}	801.6 ± 211.2	414.5 ± 82.9	–8.81	<0.0001**
Tau	233.6 ± 81.8	240.8 ± 99.3	0.25	0.80
Ptau	51.7 ± 14.1	46.2 ± 16.4	–1.15	0.26

Data are presented as means ± standard deviation. CSF, cerebrospinal fluid; Aβ_{42}, amyloid-beta isoform 42; tau, total tau; ptau, phosphorylated tau. [a]Pearson Chi-Square; **$p < 0.0001$.

RESULTS

Sample characteristics

Demographical and biological data of the CTR and PreAD-1 groups are shown in Table 1. Age ranged between 49 and 86 years, and educational level ranged between 3 and 20 years. There were no significant differences in age ($p = 0.17$) or years of education ($p = 0.49$) between the CTR and PreAD-1 groups. Gender distribution was also similar ($\chi^2 = 0.29$; $p = 0.59$) with women accounting for more than 70% in both CTR and PreAD-1. Regarding the AD biomarkers, CSF Aβ_{42} was significantly lower in the PreAD-1 group ($t(43) = -8.81$; $p < 0.001$). There were no significant group differences in the levels of CSF tau ($t(43) = 0.25$; $p = 0.80$) or ptau ($t(43) = -1.15$; $p = 0.26$). The APOE-ε4 allele was significantly more frequent in the PreAD-1 group than in CTR ($\chi^2 = 14.34$; $p < 0.001$), with a frequency of 57.1% versus 6.5% of carriers, respectively.

Descriptive characteristics of the PreAD-2 subjects (n = 4)

The PreAD-2 subjects ($n = 4$) had a mean age of 77.8 (6.9) years [range 70.9–86.1], compared to a mean age of 67.8 years in the PreAD-1 [range 58.2–78.3]. Their mean length of education was 11 years [range 3–18], compared to 10.8 [range 6–20] in PreAD-1 group. In PreAD-2, 50% were women, compared to 79% in PreAD-1. APOE-4 carriers represented 25% of the PreAD-2 sample compared to 55.5% in the PreAD-1. Regarding the biological data, PreAD-2 group had a mean CSF Aβ_{42} of 341.6 (124.1) pg/ml [228.5–512.2], CSF tau of 486 (128) pg/ml [389–666], and CSF ptau of 89.8 (17.4) pg/ml [75.2–114], compared to 414.5 (82.9) pg/ml, 240.8 (99.3) pg/ml, and 46.2 (16.4) pg/ml in the PreAD-1 group, respectively.

AFE-T performance in PreAD-1

Initial learning phase
The ANCOVA on correct spontaneous naming responses showed a significant main effect for run in the whole sample ($F(13,533) = 5.3$; $p < 0.001$), indicating an overall increase of naming performance throughout the learning sessions in both groups (see Fig. 3). Furthermore, the run x group interaction term was significant ($F(13,533) = 4.7$; $p < 0.001$), reflecting the steeper learning curve of the CTR group and their better overall performance ($F(1,41) = 6.9$; $p < 0.01$) (see Fig. 3). Specifically, the CTR group showed a mean learning progression of 16.7 (4.9) items ($t(30) = 18.7$; $p < 0.001$) (i.e., the difference between the first and the last learning run scores), whereas the PreAD-1 group showed a mean value of 12.1 (6.0) items ($t(13) = 7.52$; $p < 0.001$). The main statistical comparisons are shown in Table 2. When looking at the scores for each run, ANCOVA revealed significant between-group differences in runs 1, 2, 6, and 7 on the first learning day and in all the runs of the second learning day (see Fig. 3).

Immediate cued recall
In the immediate cued-recall performed at the end of the second learning day, the mean for the CTR group was 21.6 (2.4) points, and the PreAD-1 group a mean of 17.3 (4.5) points. The ANCOVA showed that this group difference was statistically significant ($F(1,40) = 15.4$; $p < 0.001$; see Fig. 3).

Forgetting rate (one-week, three-month, and six-month delayed sessions)
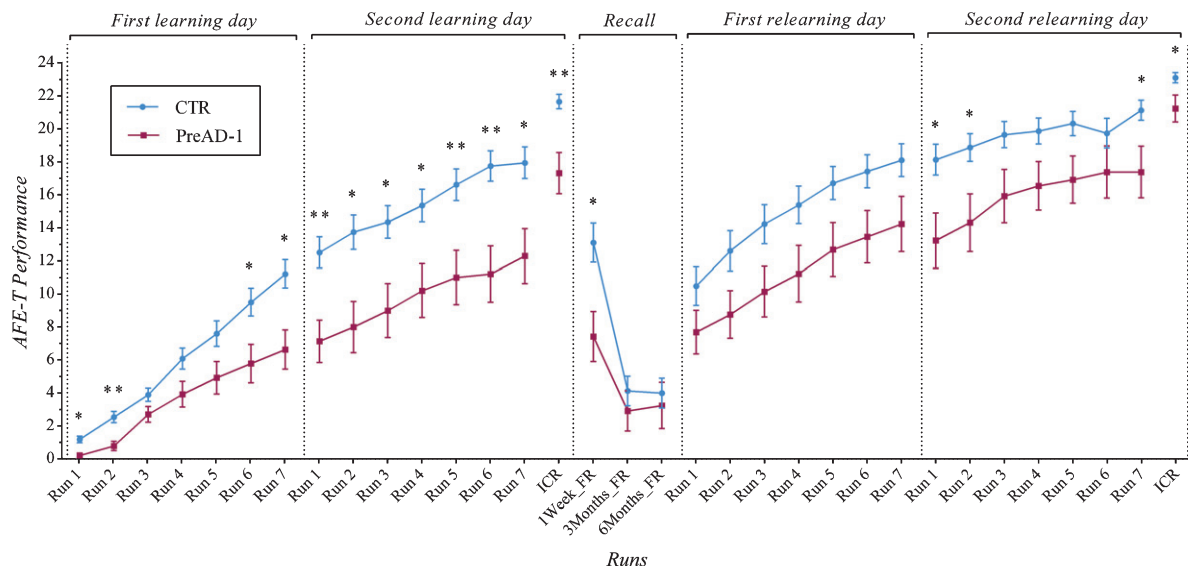Forgetting rates (one minus the ratio between each delayed session score and the score obtained on the

Fig. 3. AFE-T learning, free recall, and relearning scores of CTR and PreAD-1 groups. ICR, immediate cued recall; FR, free recall. Group differences in each point are indicated by an asterisk (*$p < 0.05$, **$p < 0.01$).

Table 2
ANCOVA of learning, recall and relearning AFE-T scores between CTR and PreAD-1 group (only the first and last run of each learning and relearning session are shown)

	CTR ($n = 31$)	PreAD-1 ($n = 14$)	F	p value
LS1_R1	1.2 ± 1.1	0.2 ± 0.6	7.39	0.01*
LS1_R7	11.2 ± 4.8	6.6 ± 4.4	6.73	0.01*
LS2_R1	12.5 ± 5.2	7.1 ± 4.8	7.80	<0.01**
LS2_R7	17.9 ± 5.4	12.3 ± 6.2	6.99	0.01*
ICR	21.6 ± 2.4	17.3 ± 4.5	15.37	<0.01**
1 Week_FFR	0.27 ± 0.3	0.47 ± 0.3	3.75	0.06
1 Week_CFR	0.16 ± 0.1	0.29 ± 0.2	4.09	0.05
3 Months_FFR	0.77 ± 0.2	0.83 ± 0.2	0.23	0.64
3 Months_CFR	0.46 ± 0.2	0.66 ± 0.2	4.83	0.03*
6 Months_FFR	0.79 ± 0.2	0.81 ± 0.3	0.01	0.99
6 Month_CFR	0.50 ± 0.2	0.61 ± 0.3	0.35	0.55
RLS1_R1	10.5 ± 6.5	7.7 ± 4.8	0.79	0.38
RLS1_R7	18.1 ± 5.1	14.2 ± 6.0	2.93	0.10
RLS2_R1	18.1 ± 5.2	13.2 ± 6.0	5.32	0.03*
RLS2_R7	21.1 ± 3.4	17.4 ± 5.6	5.35	0.03*
RL_ICR	23.1 ± 1.7	21.2 ± 2.9	5.21	0.03*

Data are presented as means ± standard deviation. LS1, 1st learning day; LS2, 2nd learning day; R, run number; ICR, immediate cued recall; FFR, free forgetting rate; CFR, cued forgetting rate; RL, relearning. *$p < 0.05$; **$p < 0.01$.

last learning run) were analyzed in each group. In this way, the forgetting rate represents the mean percentage of object/name items previously learned that were forgotten. At the one-week free recall session, the CTR had a forgetting rate of 0.27 (0.3) while the PreAD-1 group had a forgetting rate of 0.47 (0.3). ANCOVA revealed that this difference was not statistically significant ($F(1,36) = 3.7$; $p = 0.06$). When including the phonemic cue, the forgetting rate

decreased to 0.16 (0.1) in the CTR and 0.29 (0.2) in the PreAD-1; the difference showed a marginal significant trend ($F(1,36) = 4.1$; $p = 0.05$). At three months, the CTR group showed a forgetting rate of 0.77 (0.2) in the free recall and the PreAD-1 showed a forgetting rate of 0.83 (0.2). When the cue was administered, the CTR got a forgetting rate of 0.46 (0.2) whereas the PreAD-1 decreased only to 0.66 (0.2). The latter difference was statistically significant ($F(1,36) = 4.8$; $p = 0.03$). At six months, both CTR and PreAD-1 once again had similar forgetting rates in the free recall: 0.79 (0.2) and 0.81 (0.3), respectively. When including the cues, the between-group differences did not reach statistical significance ($F(1,36) = 0.4$; $p = 0.55$), with forgetting rates of 0.50 (02) and 0.61 (0.3) in the CTR and PreAD-1 groups, respectively. Both groups benefited significantly from the phonemic cues in each of the delayed sessions (see Supplementary Figure 1).

Delayed recognition

Recognition scores were based on correctly recognized objects and correctly rejected foils (24 for each, total score = 48). The ANCOVA showed a significant group difference in the one-week delayed visual recognition scores ($F(1,41) = 5.5$; $p = 0.02$), although the difference between the means was only 0.5 points (CTR = 47.9 ± 0.4, PreAD-1 = 47.4 ± 1.4). No significant group differences in visual recognition were observed at three months or six months. Neither did the picture-word matching task at six-month post

learning reveal any group difference in recognition memory.

Relearning phase

The ANCOVA on correct spontaneous naming responses showed a significant main effect of run ($F(13,533) = 11.7$; $p < 0.001$) in the whole sample, reflecting an overall significant increase of naming performance throughout the relearning sessions in both groups (see Fig. 3). However, an contrary to the learning phase, similar steepness of the relearning curves between the CTR and the PreAD-1 groups were found (Time x Group interaction, $F(13,533) = 0.7$; $p = 0.75$). The results showed significant overall differences in performance across groups ($F(1,41) = 4.4$; $p < 0.05$). When looking at the scores for each run, ANCOVA revealed significant between-group differences in runs 1, 2, and 7 on the second relearning day. Significant differences were observed by the ANCOVA on the immediate cued recall ($F(1,39) = 5.2$; $p = 0.03$), with the CTR scoring 23.1 (1.7) compared to 21.2 (2.9) in the PreAD-1 group.
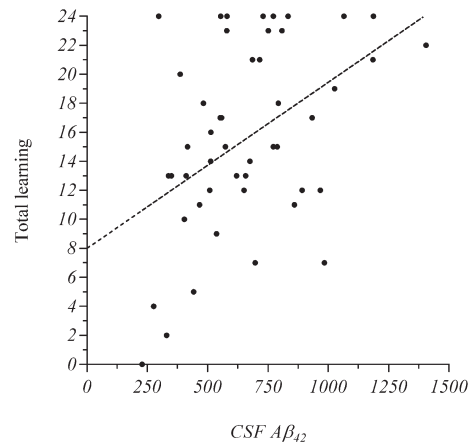
Descriptive analyses of the AFE-T performance in PreAD-2

The PreAD-2 group learned a total of 7.8 (6.9) object/name pairs on the AFE-T, with a range between 0 and 14 on the last learning run (see Fig. 4). This was below the mean in the PreAD-1
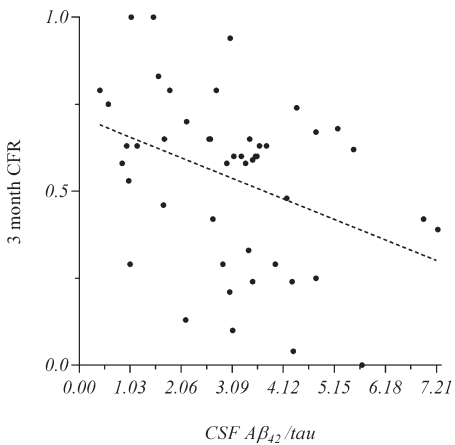

Correlation between total learning and CSF Aβ₄₂/tau ratio
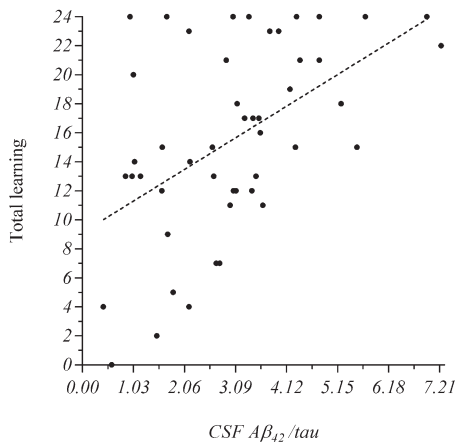
Correlation between total learning and CSF Aβ₄₂ levels

Correlation between 3 month CFR and CSF Aβ₄₂/tau ratio
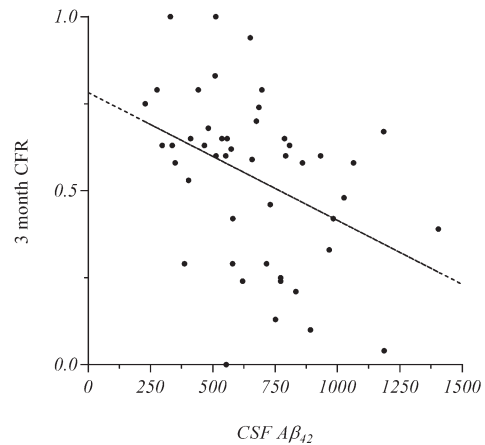
Correlation between 3 month CFR and CSF Aβ₄₂ levels

Fig. 4. Correlations between total learning and 3 month forgetting rate scores of the AFE-T and CSF Aβ₄₂ and CSF Aβ₄₂/tau ratio in the whole sample. CFR, cued forgetting rate.

group (12.3), while the standard deviation was similar in the two groups (6.2 and 6.9 for PreAD-1 and PreAD-2, respectively). As regards the forgetting rates, at the one-week free recall the PreAD-2 subjects forgot 0.4 (0.2) of the acquired items, compared to 0.5 (0.3) in the PreAD-1 group. When the cue was administered, the PreAD-2 rate remained at 0.4 (0.3) compared to 0.3 (0.2) in the PreAD-1. At three months, the PreAD-2 subjects had a mean index of 1 (0) in free-recall, decreasing to 0.8 (0.2) when the cue was included. The PreAD-1 group, in comparison, had scores of 0.8 (0.2) in the free forgetting rate and 0.7 (0.2) in the cued forgetting rate.

CSF $A\beta_{42}$ levels and AFE-T scores in the whole sample

Total learning was defined as the score obtained on the last run of the learning sessions. A significant positive correlation was found between the total learning score and CSF $A\beta_{42}$ ($r = 0.37$; $p = 0.01$). To explore the relationship of CSF $A\beta_{42}$ and recall, the correlation between the biomarker and the three-month cued forgetting rate was calculated, showing a negative correlation ($r = -0.34$; $p = 0.03$; see scatter plots in Fig. 4).

We also measured the association between the AFE-T and the ratio $A\beta_{42}$/tau. Results showed a significant positive correlation between the total learning score and the ratio ($r = 0.52$; $p < 0.001$). Regarding the three-month cued forgetting rate, a significant negative correlation was found with $A\beta_{42}$/tau ($r = -0.38$; $p < 0.01$; Fig. 4).

Standard neuropsychological tests in PreAD-1

There was no significant difference in global cognition between the CTR and the PreAD-1 group ($t(43) = -0.2$; $p = 0.9$), as assessed by the MMSE [34]. Nor was there a significant difference on the verbal intelligence score ($t(43) = 0.6$; $p = 0.5$). Crucially to the present research, no single test of the standard neuropsychological battery showed significant differences between the groups, with p values ranging from 0.07 to 0.95 (see Table 3). Regarding the word and pseudo-word span tasks included in the AFE-T (see Supplementary Table 1), significant group differences were found for the pseudo-word span ($F(1, 38) = 7.7$; $p < 0.01$) with better performance in the PreAD-1 group than in the CTR. No significant differences were observed in the word span task ($F(1, 38) = 0.1$; $p = 0.7$). Finally, the mean scores on the

Table 3
T-test of the standard neuropsychological scores between CTR and PreAD-1 groups

Parameters	CTR (n = 31)	PreAD-1 (n = 14)	t	p value
Global Cognition				
MMSE[a]	28.1 ± 1.6	28.1 ± 1.6	–0.15	0.88
WAT[a]	24.9 ± 4.6	23.9 ± 4.7	0.63	0.53
Memory				
FCSRT – FR	12.4 ± 2.8	11.1 ± 2.6	1.42	0.16
FCSRT – TR	12.5 ± 2.7	12.2 ± 3.0	0.37	0.71
FCSRT – DFR	12.5 ± 2.7	11.6 ± 3.0	0.97	0.34
FCSRT – DTR	13.7 ± 4.3	12.6 ± 4.0	0.81	0.42
Language				
BNT	11.5 ± 1.8	11.1 ± 1.9	0.66	0.51
Sem-Flu	11.0 ± 2.0	10.2 ± 2.9	1.06	0.27
Perception				
VOSP – Numbers	13.1 ± 4.6	13.0 ± 4.9	0.06	0.95
Executive functions				
TMT-A	11.1 ± 2.2	10.5 ± 2.6	0.80	0.43
TMT-B	10.3 ± 2.3	9.4 ± 2.5	1.11	0.27
Stroop-W	11.3 ± 2.0	10.7 ± 1.6	0.97	0.34
Stroop-C	11.1 ± 1.8	9.9 ± 2.2	1.76	0.09
Stroop-CW	11.7 ± 2.3	10.3 ± 1.8	1.91	0.07
SDMT	12.2 ± 2.8	11.1 ± 3.7	0.99	0.33
Digits-F	10.8 ± 2.6	11.6 ± 2.8	–0.99	0.33
Digits-B	12.4 ± 2.1	12.0 ± 2.7	0.54	0.59

Data correspond to scaled scores of the standard neuropsychological tests and are presented as means ± standard deviation. MMSE, Mini-Mental State Examination; WAT, Word Accentuation Test; FCSRT-FR, Free and cued selective reminding test free recall; FCSRT-TR, total recall; FCSRT-DFR, delayed free recall; FCSRT-DTR, delayed total recall; BNT, Boston naming test; Sem-Flu, Semantic fluency; VOSP, visual object and space perception battery; TMT-A, Trail making test A; TMT-B, Trail making test B; Stroop-W, Stroop test words; Stroop-C, Stroop test color; Stroop-I, Stroop test color-word; SDMT, Symbol digit modality test; Digits-F, WAIS Span-digit forward; Digits-B, WAIS span-digit backward. [a]Raw scores.

standard neuropsychological tests of PreAD-1 group and the descriptive scores of the PreAD-2 subjects showed that the four PreAD-2 participants had higher scaled scores on all the tests administered.

DISCUSSION

This study searched for evidence for subtle learning and/or recall difficulties in Pre-AD by employing a highly demanding associative word learning test, the AFE-T. The test had to be particularly sensitive as these cognitive difficulties are too mild to be detected by standard neuropsychological tests. Moreover, we explored the possible associations between learning and memory performance and CSF proteins in Pre-AD subjects. The results observed were very conclusive in showing initial learning difficulties in our PreAD-1 subjects, while their long-term

recall and relearning were relatively preserved. Additionally, we found that CSF Aβ_{42} levels correlated significantly with the total learning score. Our findings suggest that the AFE-T is a promising tool for characterizing the cognitive profile of PreAD-1, being sensitive enough for detecting incipient episodic memory difficulties in this population.

The usefulness of the NIA-AA staging has been demonstrated in recent reports involving Pre-AD subjects [3–6]. Mormino et al. [4] studied 166 cognitively normal individuals divided into preclinical groups for a mean of 2.1 years. PreAD-1, control, and SNAP showed improvement in performance over time (due to task repetition effects) while PreAD-2 subjects declined, suggesting that the co-occurrence of Aβ deposition and neurodegeneration (i.e., PreAD-2) accelerates cognitive decline in cognitively healthy individuals. In a recent study by Soldan et al. [5], 222 cognitively healthy subjects were followed up for a mean of 11.0 years and classified into preclinical stages. Only PreAD-2 subjects showed worse cognitive performance both at baseline and longitudinally compared with the other biomarker groups, whereas controls, PreAD-1 and SNAP groups did not differ. The authors concluded that baseline and longitudinal cognitive decline is only detected in PreAD-2 subjects. However, it should be borne in mind that these studies used standard neuropsychological tests and memory composites to evaluate the cognitive performance of Pre-AD subjects. Otherwise, Papp et al. [6] studied 260 clinically normal older adults grouped in preclinical stages using a highly demanding associative memory test (the MCT). The authors found decrements in PreAD-1 subjects' free recall score when compared with Controls.

The present study investigated the cognitive performance of PreAD-1 subjects using a highly demanding associative memory test. Unlike most memory tests used to assess Pre-AD, the AFE-T requires learning, binding and storing novel information. Forming new associations or binding unrelated information with previous semantic knowledge is thought to set high demands on cognitive processing [12]. This kind of learning may depend on the formation of new neural connections in brain areas specifically related to the acquisition of new knowledge [13] which show incipient changes in Pre-AD, such as the hippocampus and adjacent medial temporal lobe structures [19, 20]. Probably due to these higher cognitive demands related to the associative learning, the AFE-T was able to find consistent learning difficulties in PreAD-1 subjects when compared to controls. This important

finding suggests that the AFE-T is sensitive enough to identify differences between controls and Pre-AD subjects, even at the first stage of the Pre-AD phase. Furthermore, considering the fact that the PreAD-1 group performed better in the pseudo-word span, their impaired initial learning in the AFE-T could not be due to impairments in attention or working memory, factors that are strongly linked to episodic memory [50, 51]. However, it is important to note that the idea that episodic (associative) memory is the first memory system to be affected in AD has been questioned [52]. Moreover, the fact that item-based and associative memory systems are independent also remains unclear [53].

One of the main strengths of the present study is the long-term follow-up of participants, which allowed a comprehensive assessment of delayed memory and forgetting rates. Since most of the time intervals in standard neuropsychological tests range between 20 and 30 minutes [28], the assessment of longer term (days or months) forgetting rates in Pre-AD remains a field to be explored. Our analysis involving PreAD-1 subjects and controls showed differences in their raw scores at one-week free and cued recall, and at three-month cued recall. Nevertheless, these group differences were influenced by the initial learning performance, since analysis of forgetting measures showed that the PreAD-1 subjects and controls presented similar forgetting rates. These findings suggest that the initial consequences of amyloid deposition affect initial learning and encoding processes more than posterior recall and retrieval processes. Only the three-month cued forgetting rate showed significant differences between PreAD-1 subjects and controls. Though weak, the greater benefit from the phonological cueing in the CTR group than in the PreAD-1 suggests that the poorer performance exhibited by the PreAD-1 group in this long-term recall session should not be attributed merely to a "tip-of-the-tongue" effect, but to a subtle information loss. Similar results were presented in a previous study using the AFE paradigm in which MCI patients benefited less from phonological cueing than controls [26]. Regarding the secondary analysis, the small group of PreAD-2 subjects presented a similar performance in the free recall but a lower benefit from the cue. These results are in agreement with a recent report which indicated that while PreAD-1 subjects showed subtle reductions in free recall, a decline in the cued recall may represent progression to PreAD-2 stages [46]. Albeit collection of long-term forgetting rates with free and cued recall is cumbersome, these

findings suggest that they can provide valuable information for identifying memory difficulties in Pre-AD.

Another innovative memory assessment included in the AFE-T was Relearning, which was included to investigate previous learning influence. Information that cannot be remembered in a free or cued recall or recognition tests can be reactivated and detected by relearning tasks [31–33]. Relearning in the Pre-AD phase has not previously been studied. After the floor effect present in both groups at the 6-month recall session, we further explored whether there were existing but inaccessible memory traces that could be reactivated during relearning. As shown by the similar positive relearning curves, both groups were able to benefit from this intervention. Interestingly, between-group differences during this relearning phase were lower than those observed during the initial learning phase. Again, these results might suggest that initial learning is the most powerful cognitive feature for discriminating PreAD-1 subjects from controls. Clinically, this makes the testing paradigm more viable as delayed testing could be avoided. Regarding the usefulness of standard neuropsychological tests, several reports in recent years have failed to find group differences between normal aging and clinically normal at-risk subjects [8–11]. In agreement with these results, and as expected considering the inclusion criteria of the present study, we did not find any significant cross-sectional difference between our two groups in a comprehensive battery of standard neuropsychological tests. Thus, currently available standard neuropsychological tests do not seem to have sufficient sensitivity to differentiate cognitively healthy individuals with decreased CSF $A\beta_{42}$ levels from controls [12, 54].

In the present study, we also examined the association between CSF $A\beta_{42}$ levels and AFE-T performance in the PreAD-1 group. The relationship between amyloid and cognition in Pre-AD has been studied in recent years, but most cross-sectional studies have not found a relationship between amyloid levels and memory performance using standard neuropsychological tests [8, 55–58]. Only longitudinal studies have shown stronger associations between amyloid burden and future memory impairment, indicating that amyloid burden in cognitively normal individuals precedes cognitive impairment and is associated with a higher risk of future cognitive deterioration [59, 60]. In the present study, the highest statistically significant association between the CSF $A\beta_{42}$ levels and the AFE-T performance was found in the total learning score. This finding supports the view

that episodic memory decline is more closely related to amyloid levels [60], and that this link may be seen only with a highly demanding associative learning task. In line with this, an association between amyloid accumulation in the frontal cortex and cognitive performance was described in a previous study in which a highly challenging face-name associative test was administered, indicating that in addition to the medial temporal lobe and related structures, frontal regions are also critical in associative memory encoding and recall [12]. This also concurs with a previous PET study which showed that the AFE paradigm engages executive and attentional functions [23].

One notable limitation of the present study is its small sample size, limiting the power of the statistical analyses. For instance, although no standard neuropsychological test showed significant differences between CTR and pre-AD, the probability of type II error in these analyses was high. Regarding AFE-T, the comprehensive long-term assessment procedure of the AFE-T can be considered to safeguard against spurious results that may hamper these kinds of studies. With regards to the delayed recall of the AFE-T, it may appear surprising that both the CTR and PreAD-1 group showed poor performance after six months of learning. There is probably a tradeoff between learning runs and the length of maintenance, and a shorter interval might have shown a difference between the groups. Previous studies in MCI patients [26] used more training days, and those participants had better maintenance of information at six months. Another potential limitation concerns the multiple comparisons problem that arises from the large number of statistical comparisons performed. This was dealt with *post-hoc* Bonferroni corrections, albeit this is an admittedly conservative method. Lastly, although the AFE-T as a whole is not fully suitable for use in the clinical setting, it allows characterization of the different processes involved in learning and memory function in the preclinical phase of AD. In the light of the present results which identify initial learning as the most sensitive area for detecting cognitive difficulties in the PreAD-1, we are now designing and validating a shortened version of the AFE-T for use in the clinical setting.

In conclusion, using a new, highly demanding, comprehensive associative memory test, we identified significant incipient learning difficulties together with a relative preservation of the recall processes in PreAD-1 subjects. Our findings suggest that the AFE-T is a promising tool for characterizing the cognitive profile of PreAD-1 and that it is sensitive

enough to detect incipient episodic memory difficulties in this population.

ACKNOWLEDGMENTS

Authors' disclosures available online (http://j-alz.com/manuscript-disclosures/16-1173r1).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-161173.

REFERENCES

[1] 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 MV, 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.

[2] Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, Bakardjian H, Benali H, Bertram L, Blennow K, Broich K, Cavedo E, Crutch S, Dartigues JF, Duyckaerts C, Epelbaum S, Frisoni GB, Gauthier S, Genthon 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.

[3] Knopman DS, Jack CR, Wiste HJ, Weigand SD, Vemuri P, Lowe V, Kantarci K, Gunter JL, Senjem ML, Ivnik RJ, Roberts RO, Boeve BF, Petersen RC (2012) Short-term clinical outcomes for stages of NIA-AA preclinical Alzheimer disease. *Neurology* **78**, 1576-1582.

[4] Mormino EC, Betensky RA, Hedden T, Schultz AP, Amariglio RE, Rentz DM, Johnson KA, Sperling RA (2014) Synergistic effect of β-amyloid and neurodegeneration on cognitive decline in clinically normal individuals. *JAMA Neurol* **71**, 1379-1385.

[5] Soldan A, Pettigrew C, Cai Q, Wang M-C, Moghekar AR, O'Brien RJ, Selnes OA, Albert MS, BIOCARD Research Team (2016) Hypothetical preclinical Alzheimer disease groups and longitudinal cognitive change. *JAMA Neurol* **73**, 698-705.

[6] 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.

[7] Lim YY, Ellis KA, Ames D, Darby D, Harrington K, Martins RN, Masters CL, Rowe C, Savage G, Szoeke C, Villemagne VL, Maruff P (2013) Aβ amyloid, cognition, and APOE genotype in healthy older adults. *Alzheimers Dement* **9**, 538-545.

[8] Aizenstein HJ, Nebes RD, Saxton JA, Price JC, Mathis CA, Tsopelas ND, Ziolko SK, James JA, Snitz BE, Houck PR, Bi W, Cohen AD, Lopresti BJ, DeKosky ST, Halligan EM, Klunk WE (2008) Frequent amyloid deposition without significant cognitive impairment among the elderly. *Arch Neurol* **65**, 1509-1517.

[9] Mormino EC, Kluth JT, Madison CM, Rabinovici GD, Baker SL, Miller BL, Koeppe RA, Mathis CA, Weiner MW, Jagust WJ, Alzheimer's Disease Neuroimaging Initiative (2009) Episodic memory loss is related to hippocampal-mediated beta-amyloid deposition in elderly subjects. *Brain* **132**, 1310-1323.

[10] Storandt M, Mintun MA, Head D, Morris JC (2009) Cognitive decline and brain volume loss as signatures of cerebral amyloid-beta peptide deposition identified with Pittsburgh compound B: Cognitive decline associated with Abeta deposition. *Arch Neurol* **66**, 1476-1481.

[11] Villemagne V, Pike K, Chetelat G, Ellis KA, Mulligan RS, Bourgeat P, Ackermann U, Jones G, Szoeke C, Salvado O, Martins R, O'Keefe G, Mathis CA, Klunk WE, Ames D, Masters CL, Rowe CC (2011) Longitudinal assessment of AB and cognition in aging and Alzheimer disease. *Ann Neurol* **69**, 181-192.

[12] Rentz DM, Amariglio RE, Becker JA, Frey M, Olson LE, Frishe K, Carmasin J, Maye JE, Johnson KA, Sperling RA (2011) Face-name associative memory performance is related to amyloid burden in normal elderly. *Neuropsychologia* **49**, 2776-2783.

[13] Laine M, Salmelin R (2010) Neurocognition of new word learning in the native tongue: Lessons from the ancient farming equipment paradigm. *Lang Learn* **60**, 25-44.

[14] McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* **102**, 419-457.

[15] Davis MH, Di Betta AM, Macdonald MJE, Gaskell MG (2009) Learning and consolidation of novel spoken words. *J Cogn Neurosci* **21**, 803-820.

[16] Mestres-Missé A, Càmara E, Rodriguez-Fornells A, Rotte M, Münte TF (2008) Functional neuroanatomy of meaning acquisition from context. *J Cogn Neurosci* **20**, 2153-2166.

[17] Rodríguez-Fornells A, Cunillera T, Mestres-Missé A, de Diego-Balaguer R (2009) Neurophysiological mechanisms involved in language learning in adults. *Philos Trans R Soc Lond B Biol Sci* **364**, 3711-3735.

[18] Doré V, Villemagne VL, Bourgeat P, Fripp J, Acosta O, Chetélat G, Zhou L, Martins R, Ellis KA, Masters CL, Ames D, Salvado O, Rowe CC (2013) Cross-sectional and longitudinal analysis of the relationship between Aβ deposition, cortical thickness, and memory in cognitively unimpaired individuals and in Alzheimer disease. *JAMA Neurol* **70**, 903-911.

[19] Bernard C, Helmer C, Dilharreguy B, Amieva H, Auriacombe S, Dartigues J-F, Allard M, Catheline G (2014) Time course of brain volume changes in the preclinical phase of Alzheimer's disease. *Alzheimers Dement* **10**, 143-151.e1.

[20] Younes L, Albert M, Miller MI (2014) Inferring changepoint times of medial temporal lobe morphometric change in preclinical Alzheimer's disease. *Neuroimage Clin* **5**, 178-187.

[21] Tuomiranta LM, Càmara E, Froudist Walsh S, Ripollés P, Saunavaara JP, Parkkola R, Martin N, Rodríguez-Fornells A, Laine M (2014) Hidden word learning capacity through orthography in aphasia. *Cortex* **50**, 174-191.

[22] Cornelissen K, Laine M, Renvall K, Saarinen T, Martin N, Salmelin R (2004) Learning new names for new objects: Cortical effects as measured by magnetoencephalography. *Brain Lang* **89**, 617-622.

[23] Grönholm P, Rinne JO, Vorobyev V, Laine M (2005) Naming of newly learned objects: A PET activation study. *Cogn Brain Res* **25**, 359-371.

[24] Hulten A, Laaksonen H, Vihla M, Laine M, Salmelin R (2010) Modulation of brain activity after learning predicts long-term memory for words. *J Neurosci* **30**, 15160-15164.

[25] Grönholm P, Rinne JO, Vorobyev VA, Laine M (2007) Neural correlates of naming newly learned objects in MCI. *Neuropsychologia* **45**, 2355-2368.

[26] Grönholm-Nyman P, Rinne JO, Laine M (2010) Learning and forgetting new names and objects in MCI and AD. *Neuropsychologia* **48**, 1079-1088.

[27] Rabin LA, Paré N, Saykin AJ, Brown MJ, Wishart HA, Flashman LA, Santulli RB (2009) Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. *Aging, Neuropsychol Cogn* **16**, 357-376.

[28] Lezak MD, Howieson DB, Loring DW, Hannay HJ, Fisher JS (2014) *Neuropsychological Assessment*, University Press, Oxford.

[29] Elliott G, Isaac CL, Muhlert N (2014) Measuring forgetting: A critical review of accelerated long-term forgetting studies. *Cortex* **54**, 16-32.

[30] Geurts S, van der Werf SP, Kessels RP (2015) Accelerated forgetting? An evaluation on the use of long-term forgetting rates in patients with memory problems. *Front Psychol* **6**, 752.

[31] Hansen L, Umeda Y, McKinney M (2002) Savings in the relearning of second language vocabulary: The effects of time and proficiency. *Lang Learn* **52**, 653-678.

[32] MacLeod CM (1988) Forgotten but not gone: Savings for pictures and words in long-term memory. *J Exp Psychol Learn Mem Cogn* **14**, 195-212.

[33] van der Hoeven N, de Bot K (2012) Relearning in the elderly: Age-related effects on the size of savings. *Lang Learn* **62**, 42-67.

[34] 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.

[35] Van Harten AC, Smits LL, Teunissen CE, Visser PJ, Koene T, Blankenstein MA, Scheltens P, Van Der Flier WM (2013) Preclinical AD predicts decline in memory and executive functions in subjective complaints. *Neurology* **81**, 1409-1416.

[36] Valech N, Mollica MA, Olives J, Tort A, Fortea J, Lleo A, Belén S-S, Molinuevo JL, Rami L (2015) Informants' perception of subjective cognitive decline helps to discriminate preclinical Alzheimer's disease from normal aging. *J Alzheimers Dis* **48**, S87-S98.

[37] Grober E, Buschke H (1987) Genuine memory deficits in dementia. *Dev Neuropsychol* **3**, 13-36.

[38] Kaplan E, Goodglass H, Weintraub S (1983) *Boston Naming Test*, Lea & Febiger, Philadelphia.

[39] Goodglass H, Kaplan E, Barresi B (2000) *Boston Diagnostic Aphasia Examination-III Edition*, Pearson Canada Assessment Inc., Canada.

[40] Warrington E, James M (1991) *Visual Object and Space Perception Battery (VOSP)*, Thames Valley Test Co, England.

[41] Reitan R (1994) *Trail Making Test (TMT)*, Reitan Neuropsychology Laboratory, USA.

[42] Stroop J (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* **18**, 643-662.

[43] Smith A (1982) *Symbol Digits Modalities Test*, Western Psychological Services, Los Angeles.

[44] Wechsler D (2008) *Wechsler Adult Intelligence Scale*, Pearson, USA.

[45] Gomar JJ, Ortiz-Gil J, McKenna PJ, Salvador R, Sans-Sansa B, Sarró S, Guerrero A, Pomarol-Clotet E, Burin DI, Jorge RE, Arizaga RA, Paulsen JS, Ser T, Del, Gonzalez-Montalvo JI, Martinez-Espinosa S, Delgado-Villapalos C, Bermejo F, Jastak S, Wilkinson GS, Krueger KR, Lam CS, Wilson RS, Nelson H, Nelson HE, O'Connell A, Schrauf RW, Weintraub S, Navarro E, Uttl B, Weickert TW, Goldberg TE, Gold JM, Bigelow LB, Egan MF, Weinberger DR (2011) Validation of the Word Accentuation Test (TAP) as a means of estimating premorbid IQ in Spanish speakers. *Schizophr Res* **128**, 175-176.

[46] Davis CJ, Perea M (2005) BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behav Res Methods* **37**, 665-671.

[47] Armstrong RA (2014) When to use the Bonferroni correction. *Ophthalmic Physiol Opt* **34**, 502-508.

[48] Wolk DA, Dickerson BC (2011) Fractionating verbal episodic memory in Alzheimer's disease. *Neuroimage* **54**, 1530-1539.

[49] Dubois B, Feldman HH, Jacova C, Cummings JL, DeKosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, Meguro K, O'Brien J, 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.

[50] McCarthy R, Warrington E (1990) *Cognitive neuropsychology: A clinical introduction*, Academic Press, San Diego.

[51] Squire L, Schacter D (2002) *Neuropsychology of memory*, The GUILFORD PRess, New York.

[52] Didic M, Barbeau EJ, Felician O, Tramoni E, Guedj E, Poncet M, Ceccaldi M (2011) Which memory system is impaired first in Alzheimer's disease? *J Alzheimers Dis* **27**, 11-22.

[53] Davaci L (2006) Item, context and relational episodic encoding in humans. *Curr Opin Neurobiol* **16**, 693-700.

[54] Oh H, Madison C, Haight TJ, Markley C, Jagust WJ (2012) Effects of age and β-amyloid on cognitive changes in normal elderly people. *Neurobiol Aging* **33**, 2746-2755.

[55] Bourgeat P, Chételat G, Villemagne VL, Fripp J, Raniga P, Pike K, Acosta O, Szoeke C, Ourselin S, Ames D, Ellis KA, Martins RN, Masters CL, Rowe CC, Salvado O (2010) Beta-amyloid burden in the temporal neocortex is related to hippocampal atrophy in elderly subjects without dementia. *Neurology* **74**, 121-127.

[56] Hedden T, Van Dijk KRA, Becker JA, Mehta A, Sperling RA, Johnson KA, Buckner RL (2009) Disruption of functional connectivity in clinically normal older adults harboring amyloid burden. *J Neurosci* **29**, 12686-12694.

[57] Oh H, Mormino EC, Madison C, Hayenga A, Smiljic A, Jagust WJ (2011) β-Amyloid affects frontal and posterior brain networks in normal aging. *Neuroimage* **54**, 1887-1895.

[58] Rowe CC, Ellis KA, Rimajova M, Bourgeat P, Pike KE, Jones G, Fripp J, Tochon-Danguy H, Morandeau L, O'Keefe G, Price R, Raniga P, Robins P, Acosta O, Lenzo N, Szoeke C, Salvado O, Head R, Martins R, Masters CL, Ames D, Villemagne VL (2010) Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging* **31**, 1275-1283.

[59] Doraiswamy PM, Sperling RA, Coleman RE, Johnson KA, Reiman EM, Davis MD, Grundman M, Sabbagh MN, Sadowsky CH, Fleisher AS, Carpenter A, Clark CM, Joshi AD, Mintun MA, Skovronsky DM, Pontecorvo MJ (2012) Amyloid-β assessed by florbetapir F 18 PET and 18-month cognitive decline: A multicenter study. *Neurology* **79**, 1636-1644.

[60] Hedden T, Oh H, Younger AP, Patel TA (2013) Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology* **80**, 1341-1348.

Appendix 1.   Stimuli used in the AFE-T (novel pictures and novel words)

MIRPA	NAFE	FRUMA	RIMUTO
TREZO	SUGA	FOLUTE	GORSI
ESPAJO	AREDO	SELGO	ORIVE
UNIFA	RUPO	EMPULA	LUMBE
CROBA	PILENO	DRAPE	TECO
SAFIME	BRINE	LICA	BALIRO