

Accepted for publication in "Journal of Clinical and Experimental Neuropsychology"

Note: This is an uncorrected version of an author's manuscript accepted for publication. Copyediting, typesetting, and review of the resulting proofs will be undertaken on this manuscript before final publication. During production and prepress, errors may be discovered that could affect the content.

**Mnemonic anosognosia in Alzheimer's disease is caused by a failure to transfer on-line evaluations of performance: Evidence from memory training programmes**

Ana Rita Silva (a), Maria Salomé Pinho (a), Luís Macedo (b), Céline Souchay (c),  
Christopher Moulin (c)

a. Faculty of Psychology and Sciences of Education, University of Coimbra,  
Portugal

b. Department of Informatics Engineering, University of Coimbra, Portugal

c. Laboratoire de Psychologie & NeuroCognition (CNRS UMR 5105), Université  
Grenoble Alpes, France

Address for correspondence:

Faculdade de Psicologia e Ciências da Educação da Universidade de Coimbra

CINEICC – Centro de Investigação do Núcleo de Estudos e Intervenção Cognitivo-  
Cmporamental

Rua do Colégio Novo, s/n

3002-802 COIMBRA

PORTUGAL

ana.rita.ess@gmail.com

Abstract

**Introduction.** There is a debate about the ability of patients with Alzheimer's disease to build an up-to-date representation of their memory function, which has been termed *mnemonic anosognosia*. This form of anosognosia is typified by accurate on-line evaluations of performance, but dysfunctional or outmoded representations of function more generally.

**Method.** We tested whether people with Alzheimer's disease could adapt or change their representations of memory performance across three different six-week memory training programmes using global judgements of learning.

**Results.** We showed that whereas online assessments of performance were accurate, patients continued to make inaccurate overestimations of their memory performance. This was despite the fact that the magnitude of predictions shifted according to the memory training. That is, on some level patients showed an ability to change and retain a representation of performance over time, but it was a dysfunctional one. For the first time in the literature we were able to use an analysis using correlations to support this claim, based on a large heterogenous sample of 51 patients with Alzheimer's disease.

**Conclusion.** The results point not to a failure to retain on-line metamemory information, but rather that this information is never used or incorporated into longer term representations, supporting but refining the mnemonic anosognosia hypothesis.

KEYWORDS:

Global judgements of learning, Metacognition, Dementia, Memory training

**Mnemonic anosognosia in Alzheimer's disease is caused by a failure to transfer on-line evaluations of performance: Evidence from memory training programmes**

Metamemory is described as the awareness of one's memory abilities, including judgments and knowledge about memory (Metcalf, 2008; Shaked, Farrel, Huey, Metcalfe, Cines et al., 2014). Metamemory allows for the control and monitoring of self-initiated processes such that if someone makes inaccurate or inappropriate assessments of ongoing cognitive function, self-regulated learning will be suboptimal. Studies of metamemory processes have particular importance in effective care for patients with Alzheimer's disease (AD) because knowing how the impairment in memory is influenced by awareness, one may improve function with cognitive training and through feedback (Antoine, Nandrino, & Billiet, 2013; Clare et al., 2004, Green et al., 1993).

Metamemory can be operationalised as the relationship between subjective evaluations (predictions) and actual performance (see Ernst, Moulin, Souchay, Mograbi, & Morris, 2015 for review in Alzheimer's disease). In AD, the status of metamemory accuracy is currently under debate. Some studies point to a preservation of metamemory function (e.g. Gallo et al., 2012; Lipinska & Backman, 1996; Moulin, Perfect, & Jones, 2000; Moulin et al., 2003; Waring, Chong, Wolkon & Budson, 2008;), whereas others point to a deficit (e.g. Barret et al., 2005; McGlynn, 1991; Shaked et al., 2014; Souchay et al., 2002). One way of thinking about this apparent contradiction in the literature is to think about the basis on which people make their metamemory predictions. Souchay (2007) pointed to a fractionation of metamemory based on the cues used to make the metamemory judgements. When the judgements rely upon an updating of beliefs about memory, there is a deficit: people with AD 'forget' that they have a bad memory, even though, on-line, in the middle of a task, they show appropriate awareness.

Metamemory is often referred to in relation to another term – anosognosia – that describes the clinical manifestation of unawareness of deficit (e.g. Vogel et al., 2005). A current debate in the field considers the extent to which metamemory and anosognosia can be differentiated (e.g. see Cosentino, Metcalfe, Butterfield, & Stern, 2007; Souchay, 2007). That is, can patients be simultaneously accurate in predicting their performance, but somehow unaware of their deficit? This idea is encapsulated in the fractionation of metamemory in Alzheimer’s disease: anosognosia could arise because patients cannot build a realistic, up-to-date representation of their current level of memory function, even though they can adequately perform tests of metamemory in the here-and-now. In fact, Morris and Mograbi (2013) have offered just such an explanation of the relation between anosognosia and metamemory. They suggest that in Alzheimer’s disease there is a *mnemonic anosognosia*, based on a failure to record the outputs from online monitoring processes and update representations of self-performance. In this paper, we shall refer to these here-and-now judgements of performance, made during or shortly after conducting a memory task, ‘on-line evaluations’ in keeping with Morris and Mograbi’s terminology.

One manifestation of this deficit in updating representations of memory task performance in AD comes from the predictions of performance made before the opportunity to study. Several studies (e.g., Moulin et al., 2000) have shown that when first asked to make predictions of performance, people with AD make very inaccurate predictions, tending to overestimate their performance. In particular, initial predictions of performance, expressed as the number of items predicted that will be recalled on an upcoming test, tend not to be different from the predictions made by healthy controls (Thomas, Lee, & Balota, 2013). For many, this is *prima facie* evidence for a lack of deficit awareness in AD; predictions do not take into account that the person is experiencing memory loss. On the

other hand, if multiple tests are carried out in the same session, or participants experience at first hand their memory abilities, their on-line evaluations are accurate; patients benefit from experience in their prediction accuracy – at least in the short term.

The idea is that an initial prediction on a memory task is not an informed judgement based on knowledge of a person's abilities or the characteristics of the upcoming test. It is not possible to know in advance how difficult a test will be, and so one must rely on generalised beliefs about performance (Connor and Dunlosky, 1992). The existing data suggests that these initial estimations are inaccurate and that they therefore reflect dysfunctional expectations about memory performance and anosognosia. However, one interpretation is that people with Alzheimer's and healthy controls merely have the same generalised beliefs about performance. Although this failure to update beliefs (see below) is taken as evidence of a form of anosognosia, this has seldom been directly tested. The best test of this idea is to see if, over time, people with Alzheimer's disease can build an accurate representation of the change in their memory function, which is what we test here. The issue of whether patients can update their beliefs about memory lends itself to a longitudinal design, and in particular memory training. If we wish to measure the ability to update judgements it would be desirable to manipulate in some way their function over time. In fact we took an opportunistic approach, adding such metamemory measures to an ongoing programme of cognitive training. We expected that the initial predictions of performance should shift over time and according to the different levels of performance achieved in the memory training programmes.

Initial inaccuracy in AD has been most closely examined by Ansell and Bucks (2006) who looked at the ability to shift the initial prediction across several different tests. Ansell and Bucks tested a group of 18 AD patients. Participants were instructed to predict how

many words they would be able to recall from a 10-word list. The list was then read by the experimenter and after hearing the list participants were again asked to predict how many words they would be able to recall (post-list prediction). Then the recall test took place and this whole procedure was repeated for two more different lists. Finally, after a short delay, participants were asked again to predict the number of words they think they would be able to recall if a fourth list was given (delayed-list prediction). The AD patients were, as expected, less accurate than controls in their initial prediction, before exposure to the materials, but revised their predictions after being presented the list, becoming more accurate across lists (but never reaching a comparable level to controls).

The pattern for the final prediction made after a delay is critical to understand whether the patients update their predictions or not, testing the idea of mnemonic anosognosia. In this regard, Ansell and Bucks' data are difficult to interpret. Whereas the magnitude of the predictions made initially was significantly different between the first and third presentation of the ten-item list, the difference between the first and delayed list predictions only showed a trend ( $p=.05$ ), and the difference between the delayed list and the third list was not significant ( $p=.10$ ). The means in fact show the delayed prediction is somewhat higher than the third prediction, showing a return towards the very inaccurate first prediction, but this is difficult to interpret given the p-values reported, and the fact that, if anything, there is a trend for the judgement to be lower over the three versions of the test, and the delayed prediction is lower than the very first prediction. We aimed to address this very same issue with data over a longer delay. If there is a failure to update memory beliefs despite accurate on-line evaluations, we would expect that initial predictions, before completing a memory task should show no change. On the hand, if there is a transfer of the accurate on-line evaluations into longer-term metamemory representations, we should see significant changes over time.

In sum, the literature suggests that there are two forms of memory monitoring at play in metamemory judgments. First, initial predictions are made on the basis of generalised beliefs about memory function. The extent to which these predictions are accurate is shown in forward-looking predictions: someone who is metacognitively competent will make a prediction that relates well to subsequent performance. The second type of monitoring is made on-line, and represents the capacity to be aware of on-going mental operations. This type of monitoring is seen in the relationship between prior performance and subsequent predictions: to what extent does performance on a previous trial become incorporated into predictions on the subsequent trial? We conceived this as a backwards-looking prediction. These two types of monitoring, operationalised as metamemory judgements made at different points in a series of verbal learning trials, were measured for the first time over a long time period (six weeks).

The current study was part of a much larger programme of research looking at memory rehabilitation in AD, and we took the opportunity to measure the two types of metamemory in relation to three different types of memory training. The fact that a well matched sample underwent several different types of memory training and was seen over multiple time points gave us the possibility to examine how knowledge and beliefs about memory might change over the course of a memory intervention, and how these may affect people's beliefs about their memory, and their initial memory predictions. The comparison of these different types of memory training, whilst not carried out to directly investigate metamemory provided us with two opportunities to make further insights into the metamemory capacities of people with Alzheimer's disease. First, we anticipated differences between our memory training interventions in terms of memory function. Thus, we can ask if predictions of performance are sensitive to the factors which lead to

the improvement in memory. We should expect that a group who receives a successful intervention which significantly improves episodic memory function should also make predictions of performance which reflect that difference in performance. Second, the memory training procedures which we used (detailed below) offered us the ability to consider, as a secondary aim, the role of feedback in metamemory.

Ansell and Bucks (2006) showed that initial predictions did not retain their accuracy over a delay, but we predicted that with the use of memory training, in particular, participants would be able to retain a more accurate assessment of their memory function (see Gross et al., 2012). In particular, we predicted that participants would be able to use feedback from the memory training and frequent visits from the experimenter to update self-representations of performance. By feedback we mean that the experimenter provides the participants information concerning their performance on each task, at the end of the memory training session. The literature states that feedback is important both to improving learning in AD (Machado et al., 2009) and to increase awareness (Clare, 2004; Clare et al., 2013; Werheid, Ziegler, Klapper & Kühl, 2010). Thus, in that the three memory training types differ in their level of explicit feedback given, we should therefore see a shift in the initial first-list predictions in the memory training procedures that give more feedback to participants. The idea is that explicit feedback should improve metamemory accuracy, because it emphasises the discrepancy between perceived and actual performance in patients, and provides explicit knowledge about performance which can be incorporated into the ‘personal knowledge base’ in the cognitive awareness system (e.g. Morris & Mograbi, 2013).

## METHOD

### **Participants**



The sample consisted of 51 patients with mild AD (aged 62 to 80 years,  $M = 73.65$ ,  $SD = 5.498$ ) recruited from the Psychiatry and Neurology services of Coimbra University Hospital. Patients were recruited within 6 months of diagnosis being made according to the criteria of the NIA-AA workgroup (McKhann et al., 2011). The Addenbrooke's Cognitive Examination – Revised (Mioshi, Dawson, Mitchell, Arnold & Hodges, 2006) was used to establish a baseline of global cognitive function ( $M = 62.45$ ,  $SD = 7.71$ ). The study procedures were approved by the research ethics board of the Hospital Centre of the University of Coimbra (CHUC) (ethics approval number 4212), and all participants provided informed consent.

Insert Table 1 about here

The patients were randomly assigned to one of three groups of cognitive training (see below). These groups showed no differences in age, pre-morbid IQ and general cognitive status (see Table 1 for complete neuropsychological scores). One-way ANOVAs showed that there was no significant difference between groups for age,  $F(2, 48) = 1.475$ ,  $p = .241$ , level of education (years of formal education),  $(F(2, 48) = 2.340$ ,  $p = .109)$ , estimated premorbid IQ (TELPI results on the Portuguese version of NART (Alves, Simões, & Martins, 2012),  $F < 1$ ), and general cognitive status (ACE-R (Mioshi et al., 2006)  $F < 1$ ).

### **Design and Memory Training Procedures**

All three groups received memory training techniques which involved eleven visits from the same experimenter over a period of six weeks. The cognitive training groups were the following:

*Memo+ Group.* This program of memory training included a set of exercises to practice motivation, attention and memory. The exercises had progressive levels of difficulty (either with increasing the number of items to retain, or increasing the retention intervals) throughout the training sessions and were based on studies that showed the efficacy of those techniques to stimulate memory in people with impairment (Arkin, 2001; Cherry et al., 2010; Choi & Twamley, 2013; Dun & Clare, 2007; Fish et al., 2010; Lee et al., 2009; Netto, 2010; Serrano et al., 2010; Sohlberg et al., 2000; Winter & Hunkin, 1999; Zanetti et al., 2001). The structure of each session was the following: a) orientation questions (date, place, one of the news read in the newspaper/seen of TV in that day); b) one exercise of attention (e.g., cancellation task counting the time– draw a circle around all X found in a page full of letters); c) one exercise of episodic memory (e.g., Shopping list with 5 items – spaced retrieval technique – recall 1 min, 3 min, 5 min, 12 min); d) one exercise of implicit memory/ functional activities (e.g., to perform the actions needed to send a letter to someone in a correct order); e) feedback given by the experimenter about each exercise (describing weaker and stronger areas of performance) f) at the end of the 11 sessions there was a questionnaire of self-assessment, where the patient was asked of the perceived gains from the training. This was the only intervention where feedback was explicitly given on performance by the experimenter. We hypothesized that the Memo+ intervention, because it involved explicit feedback, should provide increased awareness by the participants that belonged to that intervention group.

*SenseCam Group.* In this cognitive training, an automatic digital camera (SenseCam) was used to capture still images from participants' daily life, which were then shown on a computer during sessions with the experimenter. For a review of SenseCam use in memory impairment see Silva et al. (2016). Numerous studies have indicated that such use of a wearable camera can increase memory function (e.g., Berry et al., 2007, 2009;

Brindley et al., 2011; Browne et al., 2011; Pauly-Takacs et al., 2011). The participant was instructed to wear the camera every day for as long as possible (from waking until the going to bed) in order to maximize the potential of the device and the number of images gathered. During the training sessions the participants reviewed their images and were asked to comment on what they saw. The experimenter did not ask questions or give feedback on the comments.

*Personal Diary Group.* In this group, participants were asked to write down their daily activities in a personal journal and to read the diary entries to the experimenter in each session. The Journal had organized sections to complete, by date, in order to facilitate filling in of the activities performed during each day, i.e. *Event description* (where the participant wrote the activity done in that day, for example, *had breakfast with my partner*); *Time* (where the participant registered the time of the day the event took place, for example, *8 a.m.*); *Place* (where for each event the participant registered where the event took place, for example, *in a coffee shop around the corner*); *People Involved* (where the participant described other people that were part of the event described, for example, *my partner Lucy*); and *Emotional Description* (where the participant was asked to describe how they felt during the event, for example, *relaxed and happy*). Two other sections were also included as aids for prospective memory: *Appointments* (to register appointments made in that day for future conclusion) and *Notes* (to register additional information such as current events seen on TV or a message given by some friend). The participant was instructed to fill the diary pages at the end of each day.

A comprehensive neuropsychological assessment was applied before assigning each participant for each memory training group, repeated immediately after the end of the training sessions and again six months after (See Silva et al., submitted). Table 1 presents

baseline scores for each group in each test of neuropsychological assessment before the intervention. There were no significant differences between groups in any measure at this point. In addition to a sociodemographic questionnaire and the ACE-R for cognitive screening, participants performed several cognitive tasks, with special focus on verbal episodic memory – using word lists (Wechsler, 1997) – to which metamemory questions were added (see details below). Participants in all groups also answered some questions at the end of their training sessions to assess the perceived gains of the training. This was done using a 4-point Likert-scale type questionnaire that included the following questions: 1) How do you judge the quality of the sessions you took part during these 6 weeks? 2) Did the training sessions meet your expectations? 3) Did these sessions allow you to find help in what you needed? 4) Has what you have learned allowed you to manage your difficulties more easily? 5) Would you advise a friend with similar difficulties to take part in this kind of training? 6) If you had the opportunity, would you be keen to take part in more of these sessions? As some of these questions are metacognitive, we analysed the results of this general questionnaire in this paper.

### **Materials and Metamemory Task Procedure**

The standard Wechsler Memory Scale – III (Wechsler, 1997, CEGOC-TEA, 2008) word list test was adapted to study metamemory judgments using global judgements of learning (gJOLs). In summary, this test uses 12 concrete unrelated words. There are four learning trials, where the experimenter reads the words in the same order and the participant is instructed to recall them immediately in any order. These trials are followed by an interference list (with a different 12 concrete and unrelated words) and by a short delayed recall task of the first list. After an interval of 20 to 25 minutes a long-delay recall is applied to the original list followed with a yes/no recognition task for the items from that list.

In the four learning trials of the Word Lists test, we added a gJOL measure, where the participants were asked, before being presented the list, to predict the number of words they will be able to recall after hearing the list. Immediately after the four trials' gJOLs and recall, the participants were additionally asked to predict how many words they would be able to recognize later (we specifically asked participants the following: "if you hear this list again later in this task, how many words you think you will be able to recognize?"). We called this prediction a judgement of recognition (JOR) <sup>1</sup>. No feedback on performance was given. We gathered these data at baseline (before the beginning of cognitive training) and after memory training (approximately one week after the last memory training session) using identical materials. All 51 patients included in the study had complete data on all predictions and recall measures.

## RESULTS

The results section is organised as follows: First the actual performance in the Word Lists test of the three groups is considered, according to trial and to visit (baseline (Visit 1) versus post training (Visit 2)). The critical issue here is whether there has been any improvement in episodic memory performance over time in the three groups. Next, the magnitude of the gJOLs is analysed, following the same design. In these analyses, we are interested primarily in whether the gJOLs show a shift according to the type of memory training, and whether we observe the initial overestimate of performance that is typical of AD. This sensitivity analysis (cf. Moulin et al., 2000) will allow us to see if the patients are aware of the effects of the memory training and the repeated learning trials on their

---

<sup>1</sup> These predictions were also analysed but not presented here. We concluded that participants did not understand the prediction of future recognition performance, and in fact predicted recall (See Moulin, 2002, for a analysis of the inefficacy of predictions of recognition in AD patients as a method to understand their metacognitive abilities).

memory. The third set of analyses will, following convention, consider the non-directional discrepancies between memory performance and predictions, with the aim of seeing whether metacognitive accuracy is altered by the memory training procedures. A subsequent analysis will consider the correlations between variables to further analyse the relative accuracy of the global predictions (cf. Connor, Dunlosky, & Herzog, 1997). We finish by analysing the results for the questionnaire of perceived gains carried out in the end of the memory training sessions for all participants, in order to assess the general beliefs in each groups of patients concerning the effects of the training in their cognition.

FIGURE 1 ABOUT HERE

### *Objective memory performance*

A 3 (Group) x 2 (Visit) x 4 (Trial) mixed ANOVA with the cognitive training group as the between subjects factor was performed. A main effect of Group was found,  $F(2, 48) = 4.18, p < .05, \eta^2_p = .167$ . Post hoc comparisons using the Fisher LSD test revealed that the MEMO+ group and SenseCam group do not differ significantly from each other ( $p = .482$ ) whereas the Diary group shows a significantly lower performance than both the MEMO+ ( $p < .01$ ) and the SenseCam ( $p = .03$ ) groups. A main effect was also found for the time of assessment,  $F(1, 48) = 23.85, p < .01, \eta^2_p = .332$ , with a superior performance in Visit 2. There was also a main effect of Trial,  $F(3, 48) = 147.06, p < .01, \eta^2_p = .754$ , with increased recall across trials, as expected. There was a significant interaction between Group and Visit,  $F(2, 48) = 20.21, p < .01, \eta^2_p = .457$ ). This is our critical analysis to see if groups' performance differed according to the memory training programme they carried out. Figure 1 shows that in the second visit the Memo + and SenseCam groups have different levels of performance from the Diary group, and moreover, the diary group shows little difference between Visit 1 and Visit 2. To examine this interaction further,

we calculated the total recall score at Visit 1 and Visit 2 and submitted these scores to a 3 (Group) x 2 (Visit) ANOVA. Unsurprisingly, there was the same interaction,  $F(2, 49) = 21.52$ ,  $p < .01$ ,  $\eta^2_p = .473$ , but planned comparison t-tests showed that total recall increased significantly for Memo+ group,  $t(16) = 4.82$ ,  $p < .01$ , and for the SenseCam group,  $t(16) = 6.17$ ,  $p < .01$ , but it decreased for the Diary group,  $t(16) = -2.81$ ,  $p < .05$ . Moreover, the difference between groups was not significant at Visit 1,  $F(2, 48) = .509$ ,  $p = .33$ , but it was significant at Visit 2,  $F(2, 48) = 10.89$ ,  $p < .01$ . No interaction effect was found for trial and visit, but a three way interaction was found between Group, Trial and Visit, ( $F(6, 41) = 4.19$ ,  $p < .01$ ,  $\eta^2_p = .149$ ), suggesting that the different forms of training across visits differently influence the rate of acquisition of items in the memory tasks.

Table 2 shows the remaining **word** list measures' scores for the three memory training groups before and after intervention, **namely recall (interference list, short and long delay) and recognition scores (hits, omissions and faults)**. **T**hese are scores for which the participants did not make predictions and are included for completeness.

Insert Table 2 about here

### ***gJOL Predictions***

Again, a 3 (Group) x 2 (Visit) x 4 (Trial) mixed ANOVA with group as the between subjects factor was performed. No main effect of group was found in the predictions,  $F(2,48)=1.65$ ,  $p=.20$ ), suggesting that, on the whole, participants in all three groups made comparable predictions of performance. A main effect was found for Visit,  $F(1,48) = 40.63$ ,  $p < .01$ ,  $\eta^2_p = .458$ , with the means indicating that participants make higher predictions on the second visit. There was also a main effect of trial,  $F(3,48) = 32.94$ ,  $p < .01$ ,  $\eta^2_p = .407$ ), where people predicted being able to recall more words in the first trial

compared to the other trials – Figure 2 shows a ‘tick’ shape curve for the predictions. There was a significant interaction between Group and Visit,  $F(2,49) = 10.36, p < .01, \eta^2_p = .302$ , indicating that groups change their predictions differently between the two visits (explored further below). There were not significant interactions between Group and Trial, nor Visit and Trial,  $F < 1$ . The three-way interaction was however significant ( $F(6,48) = 9.17, p < .01, \eta^2_p = .179$ ).

#### FIGURE 2 ABOUT HERE

In general, despite participants changing their predictions after memory training (in Visit 2) there are no significant differences between groups in how they make their predictions, even though there were differences in how the groups performed on the task. However, as has been demonstrated previously (e.g., Moulin et al., 2000), at both visits, participants made predictions on the first trial which were inappropriate. Mean predictions for the first trial ranged from 4.88 to 6.23, but mean performance was never higher than 3.64 for any group on the first trial of any test. This indicates that participants greatly overestimate their performance before starting the memory test. This can be explained by the absence of feedback in the first trial (there is no previous recall to rely on to make the prediction). Since we suggest that the predictions on the first trial have a different basis to the predictions made for the subsequent trials, we examined this pattern further by separately analysing Trial 1 in one analysis and the subsequent trials in a separate analysis.

A 3 (Group) x 2 (Visit) ANOVA examined the first trial gJOLs. No main effect of Group was found,  $F(2,48) = 1.01, p = .37$ . We did find however a main effect of Visit,  $F(1,50) = 23.02, p < .01, \eta^2_p = .32$ , where people predict being able to recall marginally more words after training ( $M = 5.86, SD = 1.20$ ) compared to before training ( $M = 5.07, SD = 1.26$ ).



No interaction was found between Group and Visit,  $F < 1$ . These results suggest that the first prediction is not influenced by which memory training group participants were assigned to. However, overall, the participants make predictions which are significantly higher after having had any form of memory training.

We performed a 3 (Group) x 2 (Visit) x 3 (Trial) mixed ANOVA with group as the between subject factor on the gJOLs for the final three trials. No main effect of Group was found,  $F(2,49) = 1.62$ ,  $p = .21$ , suggesting that the magnitude of predictions, on the whole, did not change according to the intervention used. We found a main effect of Visit,  $F(2,49) = 9.61$ ,  $p < .01$ ,  $\eta^2_p = .324$ , with participants predicting being able to recall more words in Visit 2 ( $M = 5.11$ ,  $SD = 1.15$ ) than in Visit 1 ( $M = 4.39$ ,  $SD = 1.09$ ). There was also a main effect of Trial,  $F(2,49) = 20.68$ ,  $p < .01$ ,  $\eta^2_p = .324$ , where participants increase their predictions across trials (Trial 2  $M = 4.13$ ,  $SD = 1.20$ ; Trial 3  $M = 4.55$ ,  $SD = 1.08$ ; Trial 4  $M = 4.83$ ,  $SD = 1.33$ ). We found an interaction for Visit and Group,  $F(2,48) = 13.38$ ,  $p < .01$ ,  $\eta^2_p = .358$ . No interaction for Trial and Group nor for Visit and Trial were found,  $F < 1$ . The three way interaction was significant,  $F(5,41) = 4.49$ ,  $p < .01$ ,  $\eta^2_p = .158$ . The interactions with group point to a pattern whereby predictions resemble the pattern for performance. Figure 2 shows that all predictions are the same for the first visit, but that at Visit 2, the Memo+ group and the SenseCam group predict performance which is higher than the diary group. That is, whereas the diary group continue to make predictions at the same level as in the first visit, the other two groups increase their levels of prediction. Importantly, this is the same pattern as for actual performance.

### ***Prediction accuracy***

We operationalized the accuracy of the gJOLs as the unsigned difference between the number of items predicted and the number of items correctly remembered (Hertzog et al.,

1990; Moulin et al., 2000). Figure 3 represents the accuracy curves across the four trials of recall, both before (Visit 1) and after the intervention (Visit 2), for the three intervention groups (Memo+, SenseCam and Diary). Figure 3 shows that participants are on the whole very inaccurate on the first trial, with predictions which are, on average, between 2 and 4 items different from actual performance.

### FIGURE 3 ABOUT HERE

Given this pattern, we again decided to examine accuracy separately according to Trial, with Trial 1 and Trails 2, 3 and 4 analysed separately. A 3 (Group) x 2 (Visit) mixed ANOVA was conducted for Trial 1. We found no main effects of Group nor Visit,  $F < 1$ , and also no significant interaction was found for Group and Visit,  $F < 1$ , indicating that accuracy in the first trial is not influenced by the kind of memory intervention used for improving objective performance. The mean unsigned difference between prediction and performance varied between 2.18 and 3.53. By way of comparison with other published studies, Moulin et al. (2000) found these initial predictions to have accuracy scores of between 1.94 and 4.37. Ansell and Bucks (2006) report a signed difference score of 2.75 for these first predictions. In short, there is no change in accuracy of these initial accuracy judgements according to the type of memory training procedure, nor over time. Moreover, these inaccurate overestimations of performance are in keeping with the published accuracy scores in previous studies.

A 3 (Group) x 2 (Visit) x 3 (Trial) mixed ANOVA with Group as a between subjects factor was conducted. We observed no main effects of Visit, Trial or Group in these accuracy rates,  $F < 1$ , suggesting that, as for the first trial, the accuracy for the subsequent trials is not affected by any memory intervention programme. We found one significant

interaction between Visit and Trial,  $F(2,49) = 5.59$ ,  $p < .01$ ,  $\eta^2_p = .102$ , with the means suggesting that in Visit 1 people show poorer accuracy across trials (the scores move further from 0, the perfect accuracy rate), whereas in Visit 2 participants become more accurate across trials. In short, we found no differences in the accuracy of predictions according to group, suggesting that, during the task, regardless of group, participants all have the same access to information on which to make accurate predictions.

### *Correlational analysis*

In order to examine the accuracy of the groups' predictions as a whole we analysed the correlations between the predictions (gJOLs) and recall. The rationale for this analysis is that groups which are on the whole metacognitively accurate will show reliable correlations between their predictions and performance. As an example, across several experiments, Connor et al. (1997) found a correlation of up to  $r = .66$  between predictions and performance in groups of around 40 healthy older adults. They found a pattern of correlations which changed according to when the prediction was made. Participants predicted their performance before study, between study and test and after test. Correlations were lowest (and actually negative) for the predictions made before test. Given the accuracy data above, we might expect to find little correlation between gJOLs and performance in the first trial, but which improves over trial. Our sample size of 51 permits, for the first time, this kind of analysis in Alzheimer's disease. However, because our between-subject manipulation influenced recall at the second time point, we restrict our analysis to the first visit.

Insert Table 3 about here

Table 3 shows very little relation between the predictions made before studying the list and the recall on the first trial,  $r = .037$ , in line with the idea that these initial first trial predictions do not capture recall performance. However, after this initial inaccurate prediction, as a group, people's predictions are reflective of their actual performance. There were strong correlations between performance and prediction across the subsequent trials, shown on the diagonal, with  $r$  values between .395 and .492. This analysis confirms the pattern in healthy groups by Connor et al. (1997).

One can also consider the correlations between recall in one trial and predictions in the next: this looks at how performance on the trial before is used in the subsequent prediction. Table 3 shows that all recall trials are significantly correlated with the subsequent prediction, from the first to the fourth trial (for example, Trial 1 recall correlates with the prediction on Trial 2,  $r = .58$ ,  $p < .01$ ). This brings additional evidence for the theory of online monitoring, obtained by the information participants get from previous trial's performance that contributes to an adjustment of predictions across trials.

### *Perceived gains of memory training*

A questionnaire of perceived gains was carried out with all participants. For the six questions of this questionnaire a 4-point Likert scale was used. A higher score indicated that participants generally evaluated the training sections more positively. Table 4 presents the medians and interquartile ranges for this questionnaire. We performed statistical analysis for the total score in the questionnaire. A non-parametric test was performed (Kruskal-Wallis) showing group differences,  $H(2) = 33.33$ ,  $p < .01$ . Pairwise comparisons obtained from the Kruskal-Wallis analysis revealed that the MEMO+ group and SenseCam group do not differ significantly with each other ( $p = .689$ ) whereas the Diary group shows a significant difference to both the MEMO+ ( $p < .01$ ) and the

SenseCam ( $p < .01$ ). Participants that took part in SenseCam and Memo+ training conditions perceived the sessions more positively than the Diary groups, which is in line with our finding that the magnitude of gJOLs increases according to the memory training.

Insert Table 4 about here

## DISCUSSION

This study aimed to examine two patterns of metamemory performance found in patients with AD, but for the first time in the context of memory training: first, the inaccuracy and inflexibility of initial predictions; and second, the ability of AD patients to correctly monitor their memory online. Our rationale was that memory training might lead to shifts in memory performance over time that would enable us to examine whether similar shifts occur in prediction levels, and whether memory improvement leads to increased metamemory accuracy (especially for these deficient first predictions). Secondly, we took the opportunity to measure two initial predictions with a long time period (seven weeks – six weeks of memory training plus one-week interval between the end of the training and the post intervention assessment visit) between them. We were interested in whether any gain in accuracy of predictions be maintained across this period, especially given that patients were being reminded about their memory function throughout this period (and in the Memo+ group this feedback was explicit).

Most importantly, two of our three memory training programmes were successful in their aim. For free recall, across trials, the SenseCam group (passive memory training) and the Memo+ group (paper and pencil intensive memory training with feedback) had superior performance compared to baseline and compared to the Diary group. The first critical question was whether the prediction magnitude would be sensitive to such shifts in

performance. The second critical question was whether the improvement in memory and involvement in a memory training programme might lead to more appropriate predictions of performance made before the first trial.

Focussing on the magnitude of predictions made by participants on the first trial, we found no group differences, and no interaction with group: on average, participants merely predicted higher recall on the second visit than at baseline. These mean predictions were all overestimations of performance. In short, before the first trial of the first memory test, participants overestimate their performance. After six weeks of training, the second time that they receive the test, they judge that their memory will be better than the first time. That is, they continue to overestimate their memory function. This is somewhat difficult to interpret because on the one hand, as their predictions shift upwards, so does performance (at least in two of the groups), which suggests some sort of relative metacognitive awareness. On the other hand, given their actual performance, the magnitude of these predictions, to be accurate, should shift downwards, not upwards. This is thus indicative of a metacognitive failure. The correlational analysis and results of our post-intervention questionnaire help clarify this picture. The lack of a correlation between predictions and performance suggest that these initial predictions do not access memory performance evaluations in a meaningful way. Secondly, the questionnaire findings point to a generally positive evaluation of the memory training (in the groups where it worked).

In sum, these first predictions can be taken as reflecting expectations of memory function that do not reflect idiosyncratic access to actual memory abilities, and that do not change over time. We found that these first trial evaluations were inaccurate and remained inaccurate over time. Even where we were able to improve memory function, and where

we provided feedback during the memory training programmes, we were not able to influence the accuracy of people's initial predictions. The second time participants made a first trial prediction they seemed to add a just fixed amount to the prediction they made for the first time they encountered the test. From a clinical perspective this study suggests that asking predictions of performance before performing a test is not a suitable means of gathering information about metamemory function. Although, of course, it remains to see whether these inaccurate initial predictions, whilst not being predictive of performance, may actually be related to real world memory or learning behaviours which might be of interest to clinicians.

One possible limitation is that without a healthy control group it is difficult to determine whether this lack of adaptability in people's predictions is actually normal or not. At least, using memory training paradigms, large scale research programmes have found that even two years later, healthy older adult groups retain some knowledge of memory training interventions and their effect on memory, as measured by questionnaires (e.g. Bottiroli et al., 2007). For evidence that there is a longitudinal pattern of change in Alzheimer's disease which differs from healthy controls, we can draw upon studies which have used questionnaire methods and which have assessed the relationship between self-rated change and actual cognitive (including memory) decline. Buckley et al. (2010) for instance looked at changes over an average of approximately three years in metacognition using a brief questionnaire in 535 people without dementia and 152 people with dementia. They found that amongst controls, there was the expected positive relationship between metacognitive ratings of change and actual cognitive change. In contrast, within their AD group there was a negative relationship: people who had worse cognitive performance, actually rated their performance as better. Thus it seems reasonable to assume that metacognitive evaluations differ between healthy and Alzheimer's disease groups, and that dysfunctional assessments of change characterise Alzheimer's samples. Of course,

whether such results would be replicated in healthy groups using our fine-grained approach using gJOLs is something which should be determined in future research.

From a theoretical viewpoint, at first glance, our data challenge the idea that it is a mnemonic anosognosia that is behind Alzheimer's patients' inability to update their predictions over time (Ansell and Bucks, 2006; Souchay, 2007; Morris & Mograbi, 2012). The fact that predictions shift significantly upwards over time points to the fact that a representation of changed performance has in fact been made and retained (and this is based on internal representations because we only see this shift in the two groups which actually did show some improvement according to the training given). Unfortunately, however, this representation is not accurate, and the revision is in the wrong direction for it to be judged as metacognitively accurate. Thus, whilst it appears people with Alzheimer's are able to update and retain representations about performance in the long term, they are unable to translate what they have learned during a task into more accurate first-trial predictions of performances. We might describe this as an ability to transfer on-line metamemory evaluations into meaningful generalised beliefs about an individual's memory function. Although some beliefs about memory can be updated and retained, information gained in the course of doing a task is not incorporated into the kind of long-term representations that are tapped in these initial first list predictions. Given that we find inaccurate pre-study predictions in healthy groups (Moulin et al., 2000; Connor et al., 1997) we might argue that this lack of transfer is not particular to Alzheimer's disease.

This pattern of performance is perhaps most clear in the Memo+ group, who we hypothesised should show increased metacognitive accuracy as a result of having had repeated explicit feedback as part of the memory training programme. We find no support



for this idea. Instead, this group continue to overestimate their performance, even if they do take on board the change in their memory function as a result of the training programme. Even with explicit feedback from the experimenter about performance, there is no transfer from the on-line evaluations which is seen at subsequent test sessions.

Turning to the on-line evaluations made for trials 2 to 5, where the participants have had a chance to experience one round of study and test on the word list, we find, as with previous studies, that people with AD make more accurate predictions of performance. The magnitude of the gJOLs made increased significantly across trials as memory function improved. In terms of accuracy, we found that accuracy did not change according to trial or group, and that accuracy was approximately a discrepancy of 1 item. This is in keeping with published discrepancies on similar tasks (e.g. Moulin et al., 2000, Experiment 2,  $M = 1.50$  (Alzheimer's disease),  $M = 1.68$  (Healthy older adults)). Thus our study adds to the data that suggests this form of on-line global evaluation of memory is not impaired in Alzheimer's disease. This is one area of metamemory where people with Alzheimer's disease show intact performance. Over six weeks, it was not possible to improve this level of performance, and the graph, and comparison with previous results suggests that in fact we may just show a ceiling effect for these data. This idea of accurate on-line monitoring was borne out in the correlational analysis. For the first time, we were able to show that the recall of a subsequent trial correlated with the prediction for the upcoming trial. This can be taken as showing that information from prior performance informs subsequent predictions. This is in line with published studies showing that retrospective judgements about memory performance are accurate in Alzheimer's disease (e.g. Moulin, James, Perfect & Jones, 2003).

### ***Conclusion***

We find that whereas predictions made within a task are accurate and shift according to changes in memory function brought about by memory training, the initial first-list predictions of performance do not show such accuracy. Although predictions of performance shift in line with changes in performance, i.e. patients predict that their memory performance will be improved after their memory training; this shift is in the wrong direction when we compare it to actual performance. That is, initial predictions of performance in Alzheimer's disease fail to incorporate information from on-line monitoring, and continue to be gross overestimations of performance. We suggest that this represents a failure to transfer online information into crystallised representations of subjective performance, rather than an inability to update or remember information about memory performance in the long term. Whereas representations of performance *did* change during our six-week interventions, accuracy of these initial predictions did not improve. According to the leading theory in this domain, the Cognitive Awareness Model (Morris & Mograbi, 2013), this would be due to a failure to transfer information from on-line evaluations into long term representations, which Morris and Mograbi describe as *mnemonic amnesia*. However, our data suggests that this failure to transfer may not be about memory per se, since some change in self-representation appears to be registered and retained, and not forgotten. Instead it seems that initial pre-study predictions and subsequent on-line evaluations measure different capacities, and there is very little incorporation of the latter into the former. The extent to which this is the case even in healthy populations is a priority for research into metamemory, where there are very few longitudinal studies of this nature. For clinical groups, our research explains why many patients set dysfunctional goals in rehabilitation settings: initial uninformed predictions do not take on board recent experiences of memory function. The fact that in two of our training programmes patients did show a shift in their prediction values means that it is not impossible that some rehabilitation procedure could help patients with Alzheimer's

disease incorporate feedback from performance into their metamemory judgements and retain it in the form of realistic expectations of performance.

*Acknowledgments*

This work was supported by the Portuguese Foundation for Science and Technology, to Ana Rita Silva with a doctoral studentship for the project: “Memory stimulation in Mild Alzheimer’s: the role of SenseCam in a comprehensive memory training program” [SFRH/BD/68816/2010]. Céline Souchay and Chris Moulin acknowledge gratefully the funding from the Region of Bourgogne (FABER), “Conscience et Mémoire dans la maladie d’Alzheimer” and from Fondation Médéric Alzheimer, “Nouvelles technologies et Maladie d’Alzheimer : vers une utilisation optimale de l’autographe”.

## References

- Agnew, S. K., & Morris, R. G. (1998). The heterogeneity of anosognosia for memory impairment in Alzheimer's disease: A review of the literature and a proposed model. *Aging and Mental Health*, 2, 9–15.
- Alves, L., Simões, M. R., & Martins, C. (2012). The Estimation of Premorbid Intelligence Levels among Portuguese Speakers: The Irregular Word Reading Test (TeLPI). *Archives of Clinical Neuropsychology*, 27(1), 58-68.
- Ansell, E., & Bucks, R. (2006). Mnemonic anosognosia in Alzheimer's disease: A test of Agnew and Morris (1998). *Neuropsychologia*, 44, 1095–1102.
- Antoine, P., Nandrino, J., & Billiet, C. (2013). Awareness of deficits in Alzheimer's disease patients: Analysis of performance prediction discrepancies. *Psychiatry and Clinical Neurosciences*, 67, 237–244.
- Arkin, S. (2001). Alzheimer rehabilitation by students: Interventions and outcomes. *Neuropsychological Rehabilitation*, 11, 273–317.
- Barrett A., Eslinger P., Ballentine, N., & Heilman, K. (2005). Unawareness of cognitive deficit (cognitive anosognosia) in probable AD and control subjects. *Neurology*, 64 (4), 693–699.
- Berry, E., Hampshire, A., Rowe, J., Hodges, S., Kapur, N., Watson, P., & Owen, A. (2009). The neural basis of effective memory therapy in a patient with limbic encephalitis. *Journal of Neurology, Neurosurgery and Psychiatry*, 80(11), 1202–1205. doi:10.1136/jnnp.2008.164251
- Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., & Wood, K. (2007). The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: A preliminary report. *Neuropsychological Rehabilitation*, 17(4–5), 582–601. doi:10.1080/09602010601029780
- Bottiroli, S., Cavallini, E., & Vecchi, T. (2008). Long-term effects of memory training in the elderly: A longitudinal study. *Archives of gerontology and geriatrics*, 47(2), 277-289.
- Brindley, R., Bateman, A., & Gracey, F. (2011). Exploration of use of SenseCam to support autobiographical memory retrieval within a cognitive-behavioural therapeutic intervention following acquired brain injury. *Memory*, 19(7), 745–757. doi:10.1080/09658211.2010.493893

- Browne, G., Berry, E., Kapur, N., Hodges, S., Smyth, G., Watson, P., & Wood, K. (2011). SenseCam improves memory for recent events and quality of life in a patient with memory retrieval difficulties. *Memory, 19* (7), 713–722.  
doi:10.1080/09658211.2011.614622
- Buckley, T., Norton, M. C., Deberard, M. S., Welsh-Bohmer, K. A., & Tschanz, J. T. (2010). A brief metacognition questionnaire for the elderly: comparison with cognitive performance and informant ratings the Cache County Study. *International journal of geriatric psychiatry, 25*(7), 739-747.
- Cherry, K., Walvoord, A., & Hawley, K. (2010) Spaced retrieval enhances memory for a name-face-occupation association in older adults with probable Alzheimer's disease. *The Journal of Genetic Psychology, 171*, 168-181.
- Choi, J., & Twamley, E. W. (2013). Cognitive rehabilitation therapies for Alzheimer's disease: A review of methods to improve treatment engagement and self-efficacy. *Neuropsychology Review, 23*(1), 48–62.  
<http://doi.org/10.1007/s11065-013-9227-4>
- Clare, L., Wilson, B. A., Carter, G., Roth, I., & Hodges, J. R. (2004). Awareness in early-stage Alzheimer's disease: Relationship to outcome of cognitive rehabilitation. *Journal of Clinical and Experimental Neuropsychology, 26*, 215–226.
- Clare, L.; Whitaker, C.J.; Roberts, J.L.; Nelis, S.M.; Martyr, A.; Markova, I.S.; Roth, I.; Woods, R.T.; & Morris, R.G. (2013) Memory awareness profiles differentiate mild cognitive impairment from early-stage dementia: Evidence from assessments of performance monitoring and evaluative judgement. *Dementia and Geriatric Cognitive Disorders. 35*(5-6), 266-279
- Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12*, 50–71.
- Cosentino, S., Metcalfe, J., Butterfield, B., & Stern, Y. (2007). Objective Metamemory testing captures awareness of deficit in Alzheimer's Disease. *Cortex, 43*, 1004–1019.
- Dunn, J., & Clare, L. (2007) Learning face-name associations in early-stage dementia: Comparing the effects of errorless learning and effortful processing. *Neuropsychological Rehabilitation, 17*, 735 – 754.
- Ernst, A., Moulin, C.J., Souchay, C., Mograbi, D.C., & Morris R. (2015). Anosognosia and Metacognition in Alzheimer's disease : Insights from Experimental Psychology. To appear in Dunlosky, J. *Handbook of Metacognition*.

- Fish, J., Wilson, B., & Manly, T. (2010). The assessment and rehabilitation of prospective memory problems in people with neurological disorders. A review. *Neuropsychological Rehabilitation, 20*, 161-178.
- Gallo, D., Cramer, S., Wong, J., Bennett, A. (2012). Alzheimer's disease can spare local metacognition despite global anosognosia: revisiting the confidence-accuracy relationship in episodic memory. *Neuropsychologia, 50*(9), 2356–2364.
- Green, J., Goldstein, F.C., Sirockman, B.E., & Green, R.C. (1993) Variable awareness of deficits in Alzheimer's disease. *Neuropsychiatry Neuropsychol. Behav. Neurol, 6*, 159–165.
- Gross, A. L., Parisi, J. M., Spira, A. P., Kueider, A. M., Ko, J. Y., Saczynski, J. S., & Rebok, G. W. (2012). Memory training interventions for older adults: A meta-analysis. *Aging & Mental Health, 16*(6), 722–734.  
<http://doi.org/10.1080/13607863.2012.667783>
- Hertzog, C., Dixon, R., & Hultsch, D. (1990). Relationships between metamemory, memory predictions, and memory task performance in adults. *Psychology and Aging, 5*(2), 215-227.
- Lee, S., Park, C., Jeong, J., Choe, J., Hwang, Y., & Park, C., (2009) Effects of spaced retrieval training on cognitive function in Alzheimer’s disease (AD) patients. *Archives of Gerontology and Geriatrics, 49*, 289-293.
- Lipinska, B., & Bäckman, L. (1996). Feeling-of-knowing in fact retrieval: Further evidence for preservation in early Alzheimer’s disease. *Journal of the International Neuropsychological Society, 2* (4), 350-358.
- Machado, S., et al. (2009). Alzheimer's disease and implicit memory. *Arquivos de Neuro-Psiquiatria, 67*(2a), 334-342. <https://dx.doi.org/10.1590/S0004-282X2009000200034>
- McGlynn, S. M., & Kaszniak, A. W. (1991). When metacognition fails: Impaired awareness of deficits in Alzheimer’s disease. *Journal of Cognitive Neuroscience, 3*, 183–205.
- Metcalfe, J. (2008). Evolution of metacognition. In J. Dunlosky & R. Bjork (Eds.), *Handbook of Metamemory and Memory* (pp. 29-46). New York: Psychology Press.
- Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. (2006). The Addenbrooke's Cognitive Examination revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry, 21*(11), 1078–1085. doi: 10.1002/gps.1610

- Morris, R. G., & Mograbi, D. C. (2013). Anosognosia, autobiographical memory and self knowledge in Alzheimer's disease. *Cortex*, 49 (6), 1553-1565.
- Moulin, C. (2002). Sense and sensitivity: Metacognition in Alzheimer's disease. *Applied metacognition*, 197-223.
- Moulin, C. J. A., James, N., Perfect, T. J., & Jones, R. W. (2003). Knowing what you cannot recognise: Further evidence for intact metacognition in Alzheimer's disease. *Aging, Neuropsychology and Cognition*, 10, 74–82.
- Moulin, C. J. A., Perfect, T. J., & Jones, R. W. (2000). Global predictions of memory in Alzheimer's disease: Evidence for preserved metamemory monitoring. *Aging, Neuropsychology and Cognition*, 7, 230– 244
- Netto, T. (2010) [*Working memory training in older adults*] PhD Thesis. University of São Paulo.
- Serrano, J., Latorre, J., Gatz, M., & Montanes, J. (2004). Life Review Therapy using autobiographical retrieval practice for older adults with depressive symptomatology. *Psychology and Aging*, 19, 272-277.
- Shaked, D., Farrell, M., Huey, E., Metcalfe, J., Cines, S., Karlawish, J., ... Cosentino, S. (2014). Cognitive Correlates of Metamemory in Alzheimer's Disease. *Neuropsychology*, 28(5), 695–705.
- Sohlberg, M., McLaughlin, K, Pavese, A., Heidrich, A., & Posner, M. (2000). Evaluation of attention process training and brain injury education in persons with acquired brain injury. *Journal of Clinical and Experimental Neuropsychology*, 22 (5), 656-676.
- Souchay, C., Isingrini, M., Gil, R. (2002). Alzheimer's disease and feeling-of-knowing in episodic memory. *Neuropsychologia*, 40, 2386–2396.
- Souchay, C. (2007). Metamemory in Alzheimer's disease. *Cortex*, 43, 987–1003.
- Thomas, A. K., Lee, M., & Balota, D. A. (2013). Metacognitive Monitoring and Dementia: How Intrinsic and Extrinsic Cues Influence Judgments of Learning in People with Alzheimer's disease. *Neuropsychology*, 27(4), 452–463.
- Vogel, E., McCollough, A., Machizawa, M. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, 43, 500–503.
- Waring, J. D., Chong, H., Wolk, D. A., & Budson, A. E. (2008). Preserved metamemorial ability in patients with mild Alzheimer's disease: Shifting response bias. *Brain and Cognition*, 66(1), 32–39.
- Wechsler, D. (1997). *Wechsler Memory Scale–III*. The Psychological Corporation: San Antonio, Texas.



- Werheid, K., Ziegler, M., Klapper, A., & Köhl, K. (2010). Awareness of Memory Failures and Motivation for Cognitive Training in Mild Cognitive Impairment. *Dementia and Geriatric Cognitive Disorders*, 30, 155-160.
- Winter, J. & Hunkin, N.M. (1999). Relearning in Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 14, 983-990.
- Zanetti, O., Zanieri, G., Giovanni, G. D., Vreese, L. P. D., Pezzini, A., Metitieri, T., & Trabucchi, M. (2001). Effectiveness of procedural memory stimulation in mild Alzheimer's disease patients: A controlled study. *Neuropsychological Rehabilitation*, 11, 263-272

**Table 1. Mean neuropsychological assessment battery and baseline results for the complete sample (N=51)**

Neuropsychological instruments/tests	<i>SenseCam</i>	<i>Memo+</i>	<i>Control - Diary</i>
	<i>Mean</i> <i>(SD)</i>	<i>Mean</i> <i>(SD)</i>	<i>Mean</i> <i>(SD)</i>
Addenbrooke Cognitive Examination – Revised	63.12 (7.47)	63.18 (7.67)	61.06 (8.25)
Instrument for Assessment of Functionality for Adults and Older Adults (IAFAI)	20.55 (8.64)	21.29 (11.20)	16.17 (5.63)
<b>Phonemic fluency – letter P</b>	7.47 (5.32)	7.47 (3.77)	5.88 (2.54)
<b>Semantic fluency - animals</b>	10.12 (2.75)	10.53 (3.14)	8.15 (3.48)
<b>Semantic fluency - food</b>	12.71 (4.72)	11.06 (5.76)	9.29 (3.83)
Symbol-Digit Coding (WAIS-III)	15.12 (6.33)	19.24 (11.83)	20.59 (8.69)
Digit Span (WMS-III)	10.32 (2.60)	9.47 (1.66)	9.47 (1.00)
Pyramids and Palm Trees	39.06 (4.59)	42.82 (4.46)	40.41 (5.04)
Prospective memory task	3.24 (1.30)	3.18 (1.38)	3.53 (1.13)
Rivermead Behavioural Memory Test (RBMT) route task – immediate	8.24 (1.88)	8.76 (1.56)	8.47 (0.51)
Autobiographical Memory Test (AMT) total	15.76 (4.25)	16.41 (3.72)	17.88 (5.39)

World Health Organization	108.88	102.94	98.76
Quality of Life questionnaire (WHOQOL- OLD)	(14.87)	(12.89)	(10.73)

---

Table 2. Mean results in the Word Lists tasks for three groups at the two time points

Word List tasks	Baseline			Post-intervention		
	Memo +	SenseCam	Diary	Memo +	SenseCam	Diary
Interference	2.53	2.65	2.47	3.59	3.12	2.24
List recall	(.38)	(.34)	(.34)	(1.66)	(1.32)	(1.92)
Short delay	2.12	1.88	1.82	3.24	2.06	.71*
recall	(.51)	(.25)	(.52)	(2.31)	(1.25)	(.98)
Long delay	1.65	.71	1.00	3.17	1.82	.76*
recall	(.49)	(.29)	(.40)	(2.27)	(1.47)	(1.20)
Recognition	8.24	8.64	9.29	9.64	10.47	9.0
- Hits	(.39)	(.62)	(.53)	(1.86)	(2.24)	(2.42)
Recognition	3.65	3.41	2.64	2.47	1.71	3.05
- Omissions	(.35)	(.62)	(.55)	(1.87)	(2.14)	(2.38)
Recognition	2.58	3.71	2.53	1.59	2.41	1.88
- Faults	(.44)	(.47)	(.53)	(1.21)	(1.33)	(1.90)

**Notes:** \*In the data for the Word Lists’ tasks shown above, the effects found were circumscribed to the short and long delay recall tasks performance, where interaction effects were found as well as a main effect of group, indicating that the Diary group performance was significantly poorer after training comparatively to the Memo+ and SenseCam group performance (without any significant differences found between these two groups).

**Table 3. Uncorrected correlations between predictions and recall, for the complete sample of AD participants in the study, at Visit 1 (r values, n=51)**

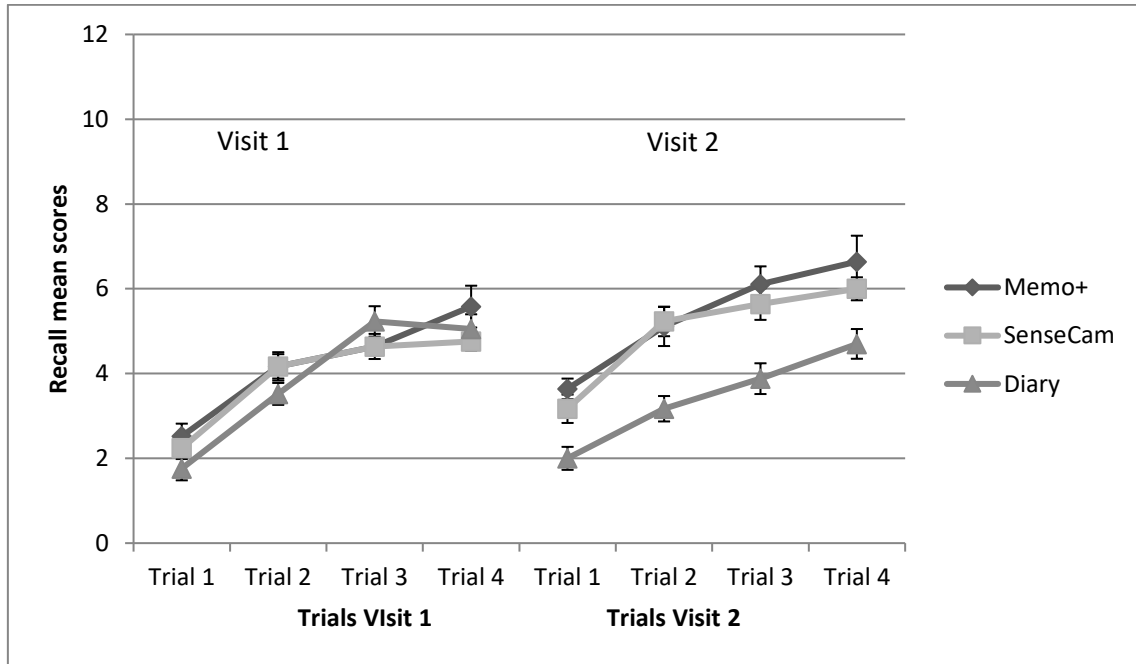
	Recall Trial 1	Recall Trial 2	Recall Trial 3	Recall Trial 4
Prediction Trial 1	.037	.002	.008	.273
Prediction Trial 2	<b>.584**</b>	<b>.460**</b>	<b>.425**</b>	.246
Prediction Trial 3	<b>.365**</b>	<b>.311*</b>	<b>.492**</b>	.176
Prediction Trial 4	<b>.491**</b>	<b>.469**</b>	<b>.425**</b>	<b>.395**</b>

\*\* . Correlation is significant at the 0.01 level (2-tailed, uncorrected).

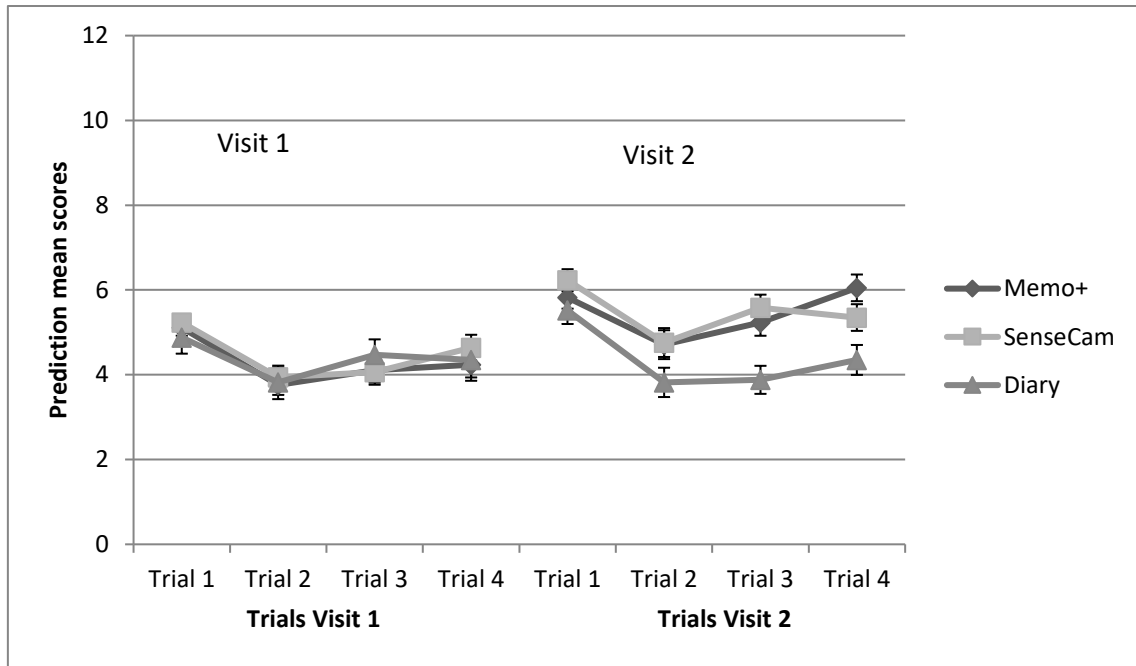
\* . Correlation is significant at the 0.05 level (2-tailed, uncorrected).

**Table 4. Medians (and interquartile ranges) on the questionnaire of perceived gains, for each cognitive training group**

	Memo+	SenseCam	Diary
1) Quality of sessions (min-1, max-4)	3.00 (.50)	3.00 (1.00)	2.00 (1.00)
2) Meet expectations (min-1, max-4)	3.00 (1.00)	3.00 (1.00)	2 (1.00)
3) Helpful (min-1, max-4)	3.00 (1.00)	3.00 (.50)	2.00 (.50)
4) Management of difficulties (min-1, max-4)	3.00 (1.00)	3.00 (1.00)	2.00 (1.00)
5) Advice others to participate (min-1, max-4)	3.00 (.00)	4.00 (.00)	2.00 (1.00)
6) Take part in future sessions (min-1, max-4)	3.00 (1.00)	3.00 (1.00)	3 (.50)
TOTAL SCORE (min-6, max-24)	19 (2.00)	20 (2.00)	14.00 (3.00)

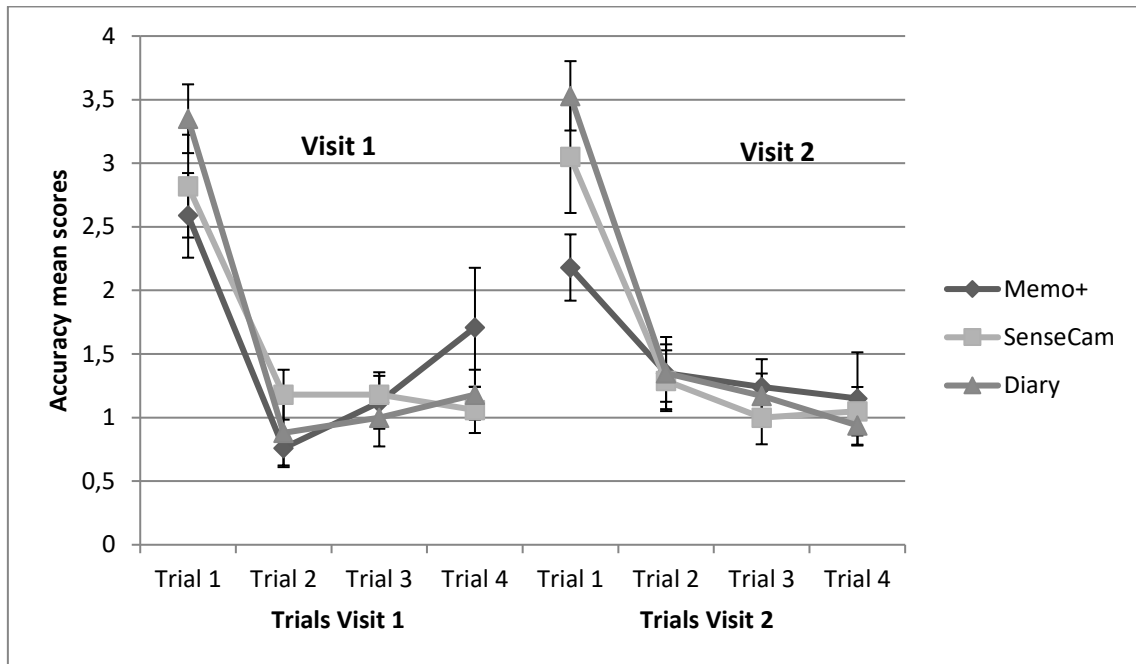


**Figure 1. Mean Memory performance by group, in Visit 1 and 2, across trials. Error bars represent 1 standard error of the mean**



**Figure 2. Mean predictions by group, in Visit 1 and 2, across trials.  
Error bars represent 1 standard error of the mean**





**Figure 3. Accuracy of metamemory predictions. Mean (and standard error) unsigned difference scores by Group, in Visit 1 and 2, and across trials**