Abstract

One of the cognitive changes associated with Alzheimer's disease is a diminution of the primacy effect, i.e., the tendency towards better recall of items studied early on a list compared to the rest. We examined whether learning and recall of primacy words predicted subsequent cognitive decline in 204 elderly subjects who were non-demented and cognitively intact when first examined. Our results show that poorer primacy performance in the Rey Auditory Verbal Learning Test delayed recall trials, but not in immediate recall trials, is an effective predictor of subsequent decline in general cognitive function. This pattern of performance can be interpreted as evidence that failure to consolidate primacy items is a marker of cognitive decline.

Keywords

Serial Position; Primacy; Memory; Cognitive Decline; Consolidation.

Introduction

A major challenge in aging research is the early discrimination between healthy and pathological aging, especially when it concerns the development of Alzheimer's disease (AD). Both healthy, cognitively intact elderly (e.g., Light, 1991) and individuals in pre-clinical stages of AD (e.g., Albert, Moss, Tanzi & Jones, 2001) present a worsening of episodic memory performance. However, although episodic memory disturbances in advanced age are not necessarily a symptom of disease or dysfunction, several studies have highlighted how subtle memory differences across cognitively intact elderly subjects can help to identify those who are more likely to convert to AD (e.g., Albert, Blacker, Moss, Tanzi & McArdle, 2007).

A recent report (La Rue, Hermann, Jones, Johnson, Ashtana & Sager, 2008; but see also Howieson et al., 2010) emphasized the role of serial position effects in free recall as a potential predictor of pathological aging. Serial position effects refer to a commonly observed pattern in free recall, where items learned early in a study list (primacy items) and items learned towards the end of the study list (recency items) are recalled better than items learned in the middle (Murdock, 1962; Glanzer, 1972). Despite some disagreements (e.g., Neath, 2010), most researchers support a dual account of primacy and recency effects, which views these effects as based on different underlying cognitive (e.g., Glanzer, 1972) and neural (e.g., Azizian & Polich, 2007; Rushby, Barry & Johnstone, 2002) mechanisms.

Lesions to the hippocampus have been associated with a conjoint reduction of the primacy effect and an intact or exaggerated recency effect (Baddeley & Warrington, 1970; Hermann et al., 1996; Milner, 1987), a pattern also observed in patients with AD (Bayley et al., 2000; Bigler, Rosa, Schultz, Hall & Harris, 1989; Carlesimo, Fadda, Sabbadini & Caltagirone, 1996; Foldi, Brickman, Schaefer & Knutelska, 2003; Gainotti, Marra, Villa, Parlato & Chiarotti, 1998),

presumably reflecting early decline of hippocampal function in this type of dementia (Jack et al. 1997, 1999; Killianny et al. 2000).

La Rue et al. (2008) compared free recall performance across two groups of healthy, cognitively intact subjects aged between 40 and 65: of these, 623 subjects had a parent who developed AD and 157 did not. Having a family history of AD is considered a strong risk factor for AD (Berti et al., *in press*; Prince, Cullen & Mann, 1994; Silverman, Ciresi, Smith, Marin & Schnaider-Beeri, 2005). La Rue et al. tested subjects' free recall performance with the Rey Auditory Verbal Learning Test (AVLT), which requires learning and recalling a list of 15 unrelated words. They defined primacy as words 1-4 in the study list and recency as words 12-15. In their results, subjects with a family history of AD did not differ from controls in measures of total recall, but showed a greater recency effect and a poorer primacy effect than their counterparts. Other known risk factors, such as APOE ε4 (e.g., Blennow, de Leon & Zetterberg, 2006; Farrer et al., 1997) and depression (e.g., Wilson et al., 2002) were not found to influence memory performance in these two groups. These results suggest that the analysis of serial position effects in free recall tasks, and of primacy effects in particular, may be a valuable tool to identify subjects who are at risk of cognitive decline or dementia.

In terms of the cognitive processes that are involved in the primacy effect, there is no clear theoretical consensus, although the most widely accepted explanation relies on the notion of increased rehearsal of early list items as compared to items learned afterwards (Rundus, 1971; Glanzer, 1972). Primacy words, being the first on the list, would therefore benefit from more opportunities for rehearsal, which in turn would lead to better encoding and, subsequently, to easier retrieval. It follows that a failure to retrieve primacy words, more so than a failure to retrieve words from any other portion of the list, may signal emerging cognitive dysfunction.

This should be especially true in delayed recall tasks, where short-lived memory processes would be less likely to be in play, and memory consolidation would be more likely to have either occurred or not (see also Gomar, Bobes-Bascaran, Conejero-Goldberg, Davies & Goldberg, 2011, for a similar argument). Consolidation (McGaugh, 2000) refers to the idea that memory traces require time and protein synthesis to become established and resistant to interference, and it is functionally dependent on the hippocampal formation, consisting of hippocampus, dentate gyrus, subiculum and entorhinal cortex (Wixted, 2004). Therefore, we hypothesize that poor retrieval of primacy items, in the absence of any symptoms of dementia, will predict increased risk of decline in cognitive function later in life; in addition, we also anticipate that poor primacy recall should be especially predictive of cognitive decline when it is observed in delayed-recall tasks, rather than in tasks of immediate memory, as it might reflect a failure to consolidate and, possibly, hippocampal dysfunction.

To test these hypotheses, we examined the level of general cognitive function longitudinally in a sample of volunteers who were cognitively intact and healthy at the preliminary visit (baseline), who then returned for up to four follow-up visits. We also tested for interactions of free recall measures with relevant factors (including APOE ε4, family history of AD, age and time from first visit) on the course of general cognitive function over time. As a measure of memory performance, we employed the AVLT (cf., LaRue et al., 2008), whereas we used the mini-mental state exam (MMSE) test (Folstein, Folstein & McHugh, 1975) as a measure of general cognitive function. The MMSE test is commonly used in both research (e.g., Yaffe et al., 2011) and clinical practice as an index of general cognitive function and cognitive decline. In addition, it is highly correlated with both other general function tests (e.g., CDR; Juva, Sulkava, Erkinjuntti, Ylikoski, Valvanne & Tilvis, 1994) and measures of brain atrophy (e.g., Frederiksen et al., 2011;

Kovacevic, Rafil & Brewer, 2009). We expected that failure to recall primacy words in the delayed trials, at the first study visit, would be predictive of cognitive decline at consecutive visits, but not failure to recall primacy words in the learning trials, which index immediate memory performance.

Methods

Subjects. A pool of 853 participants, who volunteered in the Memory Evaluation Research Initiative (MERI) and were tested at the Nathan Kline Institute, Orangeburg, NY, was available. The MERI program was established in 2003 by NP in collaboration with Rockland County, NY, and has been providing free memory and cognitive evaluations to individuals from the local community. From the total pool of 853 subjects, who participated in the MERI on different dates during the 2003-2011 time span, we extracted 204 total participants who fell within the inclusion criteria set for this study (see Table for demographic information): over 60 years of age at baseline (age range: 60- 91); at least one follow up visit; an MMSE score of 28 or higher at baseline; and no major clinical condition, psychiatric illness, or symptoms of dementia, at baseline, as established in an interview by a board-certified geriatric psychiatrist (NP). Subjects received no compensation. Each subject returned for at least one follow-up visit, typically one to two years after the previous visit, for a maximum of four follow-up visits, giving us a grand total of 625 visits, including baseline (79 subjects came for exactly one follow-up, 61 for two, 36 for three and 28 for four; in total, 204 subjects came for the baseline and visit 2 sessions, 125 subjects came for visit 3, 64 subjects came for visit 4 and 28 subjects came for visit 5; see Table). The follow up times ranged from half a year to 7 years, with mean 3.6 years and median 3.2 years. One hundred seven subjects reported a prior family history of AD (any relative) via direct questioning. In addition, 59 subjects carried the APOE ε4 allele, out of 161 subjects for

whom genotype information was available (12 ε3/ε2, 90 ε3/ε3, 50 ε3/ε4, 5 ε4/ε2 & 4 ε4/ε4). *Procedure*. After providing informed consent, subjects' vitals were examined, blood was drawn for APOE genotyping, and a general medical intake questionnaire was conducted to obtain family and medical history information, and to assess the presence of memory complaints. The MMSE score was also collected at this stage. The MMSE test requires the subject to respond to a 30 item questionnaire, addressing different cognitive areas: Orientation (time & place), Registration (i.e., short-term memory), Attention and Calculation, Recall, Language, and Copying (i.e., figure drawing). Memory was then tested as part of a neuropsychological battery, which was largely the same for all subjects, lasted ~ 2 hours, and also included psychomotor tasks (e.g., grooved pegboard), executive function tasks (e.g., digit symbol substitution test, trail making A & B), and language tasks (e.g., verbal fluency). In the AVLT, subjects are read the same list of 15 unrelated words five times and are asked to recall the words immediately after each presentation; subjects are encouraged to recall all the words on the list, including words they have already reported in previous trials. After the fifth recall trial, subjects are read an interference list (15 new unrelated words) and asked to recall this list immediately after presentation. Subsequently, subjects are required to free recall words from the original five study trials (i.e., 15 words) without hearing the word list again. Twenty minutes after the end of the last recall trial, subjects are asked again to free recall the 15 study words from the first five trials (delayed recall task). Finally, subjects perform a 2-alternative forced choice task. One of three alternative versions of this test, using different word lists for the initial presentation, the interference trial, and the new words in the 2-alternative forced-choice task, was assigned randomly to each participant. The three alternative versions were rotated longitudinally, so that each subject received a new list in each of three consecutive visits.

Study Design and Analysis. The outcome variable in our models was MMSE score at follow-up visits. The main question was whether recall performance (AVLT) of any specific portion of the 15-word list (in particular, primacy words) at baseline was especially predictive of subsequent MMSE decline. In addition, we examined whether APOE ε4 carrier status, reported family history of AD, age at baseline and time since the baseline visit affected recall performance or moderated the effect of serial position. For these purposes, we examined performance in both the five AVLT learning trials, which index immediate memory performance, and the delayed recall trial.

We used linear mixed models suitable for longitudinal data, implemented in R (R Development Core Team, 2010). In all models, MMSE scores at follow-up were modeled as a function of baseline MMSE, baseline age, time from baseline and AVLT measures were included as fixed effects; random subject effects were included to account for correlation between the repeated observations on the same subject. Since all subjects' MMSE scores at baseline were within 2 points of the maximum possible value of 30, the residuals from our models were negatively skewed and did not follow a normal distribution, as assumed by the models. Therefore, we reran the models after applying a Box-Cox transformation to the MMSE scores to make the residuals' distribution more consistent with the normality assumption. This adjustment led to negligible change in the results, which is consistent with the observation of Gelman and Hill (2007) that normality is the least important of the standard linear model assumptions. We have therefore chosen to report the results based on untransformed data, which are more readily interpretable.

We adopted a two-stage modeling strategy. The first stage entailed main effects models, in which the AVLT measure was either: a) the total number of words recalled out of 15; b) the number of words recalled from among the *first* four (primacy) words; c) the number recalled from among

the middle seven words (middle); or d) the number recalled from among the *last* four (recency) words. As the results (described later) indicated that primacy recall was a particularly strong predictor of MMSE decline, we then used primacy as our predictor of interest in the second stage.

The second stage comprised interaction models: first, a time-primacy interaction, and then exploratory models with interactions among time, primacy, and a third variable which was either ε4 carrier status, family history of AD, or age at baseline. The same two-stage modeling procedure was used with learning trial recall as the predictor of interest; in this case, each memory measure was obtained by summing over the five learning trials. The sample size for the basic model was 204 subjects with a total of 421 follow-up visits. Due to a small amount of missing information from the delayed recall and from the learning trials tests, analyses in the various alternative models included 2-4 fewer subjects. Finally, we also fit a model with both the delayed-trial and learning-trials primacy variables together to assess relative predictive power.

Results

Main effects models. The predictor of interest in the initial main effects models was either recall of all the words presented, recall of primacy words, recall of middle words, or recall of recency words for the baseline delayed trial. In each case, a significantly positive effect (better baseline recall associated with higher follow-up MMSE scores) was found. The coefficient estimates were 0.17 (Wald t-statistic = 4.98, $p<.001$) for total words, 0.51 for primacy ($t = 5.18$, $p<.001$), 0.24 ($t=3.63$, $p<.001$) for middle and 0.36 ($t = 3.37$, $p<.001$) for recency — meaning, for instance, that controlling for covariates as described above, subjects with one more primacy word recalled at baseline had MMSE scores 0.51 points higher (i.e., less decline) on average at follow up visits. These parameter estimates can be converted to effect size estimates by dividing

by 0.74, the standard deviation of the MMSE at baseline. Thus, for example, the effect size for each primacy word recalled is 0.69. These results indicate that primacy had a somewhat stronger predictive effect than middle words or recency words. Note that the issue of degrees of freedom for Wald t-statistics in linear mixed models is controversial (Verbeke & Molenberghs, 2009): since our sample was fairly large, we based all p-values on a standard normal approximation.

To examine further the importance of recall of primacy vs. other words further, we fitted two additional models: the first included both primacy and recency, while the second included both primacy and number of non-primacy words recalled. In each case, the primacy effect was notably stronger than that of the other recall variable. In the primacy vs. recency comparison, the Wald t-statistics were 4.13 [p<.001] for primacy and 1.52 [p=.13] for recency; in the primacy vs. non-primacy comparison, the Wald t-statistics were 3.30 for primacy [p<.001] vs. 1.41 [p=.16] for non-primacy. Tests of contrasts in these models did not find significant differences between the effects (e.g., a significantly larger effect of primacy than of recency), but these results do suggest that delayed recall performance at the primacy position may be relatively more predictive of MMSE change from baseline.

The results for the learning trials were quite different, as predicted. The main-effects total, primacy, middle and recency models produced coefficients 0.054 [t=4.18, p<.001], 0.12 [t=3.46, p<.001], 0.068 [t=3.49, p< .001], and 0.072 [t=1.63, p= .10], respectively, suggesting that, unlike in the delayed recall task, primacy was not a stronger predictor than middle words. In the two additional models as above, primacy was found to be relatively more predictive than recency (Wald t-statistics: 3.54 [p<.001] vs. 1.87 [p=.061] for primacy vs. recency), but not more predictive than non-primacy words in general (2.28 [p=.022] vs. 2.47 [p=.014] for primacy vs. non-primacy). The results from the learning trials, in summary, suggest that primacy recall is not

especially predictive of decline in measures of immediate memory.

To examine the relative importance of primacy recall in the delayed trial vs. primacy recall in the five learning trials, we included these two predictors in the mixed effects model simultaneously (again adjusting for baseline age and MMSE and for time since baseline). In this model, delayed recall primacy remained a significant predictor [$t=3.82$, $p<.001$], whereas primacy in the learning trials was no longer significant [$t=0.41$, $p=.680$]. However, the contrast between these terms did not yield a significant difference.

Interaction models. For the delayed trial, we found a nearly significant time-primacy interaction effect [$p=.089$] such that the speed of decline of MMSE over time was faster for subjects with poorer delayed primacy recall at baseline. Figure 1 shows model-based estimates of expected MMSE score at the 1-, 3- and 5-year marks, given baseline delayed-trial recall of 0, 2, and 4 primacy words. Inspection of the figure suggests that subjects with perfect primacy recall at baseline (all 4 primacy words recalled) are on average unchanged at follow-up; recalling only 2 primacy words is associated with cognitive decline, and recalling none is associated with faster decline.

In addition, we observed significant 3-way interactions among time, primacy, and either age at baseline [$p=.037$] or family history [$p=.001$]. These 3-way interactions indicate that the relationship between primacy recall and rate of decline described above is stronger for older subjects than for younger subjects, which is hardly surprising. More interestingly, this relationship is also stronger for subjects with a reported family history of AD than for those without.

For purposes of the above model, family history was defined as any family member with dementia. When we refitted the model with maternal history (mother with dementia) only,

instead of any family history, the three-way interaction remained significant (p=.003); in contrast, using paternal history only did not yield a significant interaction. We did not find the APOE ε4-carrier status to interact with primacy or time in its effect on MMSE course. For the AVLT learning trials, we did not observe a time-primacy interaction or a three-way interaction with age, but we did find a three-way interaction among time, primacy and family history [p=.003] in the same direction as for the delayed recall trial. In this case, the corresponding significant interaction was not observed with maternal or paternal history.

General considerations and summary. Our results suggest that MMSE decline is more strongly associated with a failure to retrieve primacy words in the delayed recall trial, rather than any of the subsequent words. In particular, the predictive power of delayed-trial primacy can be gauged by comparing coefficients of determination (R^2) for ordinary linear models predicting MMSE at follow-up. With baseline MMSE, age and time in the model, 12.6% of the variance is explained; adding delayed-trial primacy raises the explained variance to 17.9%. In contrast, primacy recall during the learning trials is not especially predictive of decline, and in general not as predictive as delayed primacy, presumably indicating that long-term memory processes may be more indicative of healthy memory function than short-term processes based on immediate memory. In addition, the effect of primacy recall appears to be stronger for subjects with a reported family history of AD.

Figure 2 presents serial position curves for the 70 “decliners” (i.e., subjects whose mean follow-up MMSE score was lower than their baseline score) and the 134 “non-decliners,” (a) in the delayed trials and (b) averaged over the five learning trials. The difference in recall between decliners and non-decliners is more pronounced in the primacy portion of the delayed-trial curve than at any other locations along the curves for either the learning or the delayed trials. This

provides visual confirmation of the conclusions based on our linear mixed effects models.

It may be noted that despite the MMSE score declines in many subjects, the mean MMSE score at visit 5 was unchanged from baseline (see Table). This reflects the fact that the 28 subjects who had five visits had a marginally significantly higher mean MMSE score at baseline than those who did not (mean = 29.43 ± 0.13 vs. 29.18 ± 0.06; Wilcoxon test $p=.09$). This also raises the possibility that our findings are impacted by systematic differences among subjects with different numbers of follow-up visits, due either to time trends in recruitment or to differential dropout. However, by conducting Kruskal-Wallis tests, we found no significant associations between number of follow-up visits and baseline age, baseline MMSE, primacy recall on the learning trials and, most importantly for us, primacy recall on the delayed trial ($p \geq .20$ in each case); in other words, we found no evidence of a relationship between follow up duration and subjects characteristics that could possibly bias the results.

Discussion

Our results confirm that the analysis of serial position effects in delayed free recall tasks may be helpful in the identification of elderly individuals who, despite being cognitively intact at baseline, are more likely to show cognitive decline over time. In our results, when testing subjects who had no dementia at baseline and an MMSE score of at least 28, poor recall of primacy words in the delayed recall trial emerged as the better predictor of decline in the MMSE score over follow-up visits, as compared to failure to recall subsequent words.

A potential criticism of our method lies in the use of the MMSE test to measure general cognitive function. The MMSE test is made up of eight separate categories, which attempt to address different cognitive areas, including memory. In fact, three of these tasks (i.e., Recall, Registration and Repetition) are eminently related to memory function. Therefore, it may be that

our outcome measure, the MMSE score, is naturally dependent on our assumed predictor, primacy recall, which in itself is memory-based. However, two main reasons made us choose MMSE over other potential alternatives. First, the MMSE test is widely used in clinical practice as a measure of cognitive function and, therefore, it is *de facto* the instrument most clinicians would adopt when evaluating cognitive performance in elderly patients. Second, we found the baseline MMSE scores to correlate significantly, in our sample, with the digit symbol substitution test score [$r=.251$; $p<.001$], which is not reliant on memory at all and has also been employed as a measure of general cognitive function in research, especially with older adults (e.g., Salthouse, 1992; Knopman, Mosley, Catellier & Coker, 2009).

Another potential issue with our results lies within the relatively small change in MMSE score reported as a function of time and primacy recall (see Figure 1), which was within a span of three points. Tombaugh (2005) has argued that, over a five-year span, as is our case, a significant change in MMSE should be of at least four points. However, a few considerations should be made. First of all, all our subjects were cognitively intact at baseline, with an initial MMSE score of at least 28; therefore, it is to be expected that, if changes are observed over a relatively short period of time, these are likely to be subtle. Secondly, and more importantly, unlike in Tombaugh's study, our subjects volunteered to return for examinations, often coming back for visits separated by just one year. Due to this, it is possible that our subjects were, on the whole, less likely to decline, at least over a limited period of time, than the subjects in Tombaugh's study, as our subjects all still displayed enough cognitive capabilities both to initiate contact for a follow up visit, and to be motivated to participate in testing. Therefore, considering that we observed significantly greater decline in general cognitive function as a function of delayed primacy recall performance, despite the high level of function of our participants, an important

clinical and empirical question to address in future studies should be whether the pattern of results that we present here can be replicated with a population of non-demented individuals whose baseline MMSE is below 27, or with subjects with mild cognitive impairment. We anticipate that, in these latter populations, the drop in performance over time for subjects with poor primacy as compared to subjects with good primacy should be more severe than with our current sample.

La Rue et al. (2008) found that subjects with a family history of AD did not differ from controls in measures of total recall, but showed poorer performance at the primacy region. Consistent with La Rue et al.'s results, we found family history to moderate the effect of delayed primacy recall on speed of cognitive decline: namely, this effect was stronger for subjects who reported a family history of AD. Importantly, in the delayed-recall trial, we found that family history moderated the effect of primacy recall on rate of decline only when originating from the mother, but not from the father. This finding is consistent with an abundance of literature indicating that elderly individuals with a maternal history of AD, rather than a paternal history of AD, show early neurological and biological signs of AD progression (e.g., Mosconi, Berti, Swerdlow, Pupi, Duara & de Leon, 2010). In addition, also consistent with La Rue et al.'s results, we did not find the APOE ε4-carrier status to interact with primacy in affecting cognitive decline.

In addition to memory consolidation, the hippocampus has been found to be also especially important for the dynamic between pattern separation and pattern completion (Yassa & Stark, 2011); pattern separation refers to the ability of discriminating between similar items, whereas pattern completion refers to merging similar items into a single representation. The aging process tends to affect pattern separation as older participants have been found to need more dissimilarity between items to be able to maintain an adequate level of discrimination. In this respect, if the

primacy advantage is due to the primacy items being inherently more distinctive than items learned later in the list - analogously to what is suggested by discrimination hypotheses of serial position effects (e.g., Brown, Neath & Chater, 2007; Murdock, 1960) - , it may be then that a decline in pattern separation could lead to a potential disappearance of the primacy effect and, therefore, signal hippocampal dysfunction. Although a pattern separation-based account of our results, unlike an account based on consolidation, does not appear to explain why in our data primacy was only a predictor of cognitive decline when tested in the delayed task, future research should consider possible links between serial position effects, the hippocampus and pattern separation.

Aging, both normal and pathological (e.g., AD), is associated, in varying degrees, to loss of episodic memory. Of clinical importance, however, is the issue of how much - or what kind of - memory loss is indicative of the latter. In our paper, we show that highly functioning, non-demented older individuals appear to show significantly greater decline in general cognitive function over time when their baseline delayed primacy recall performance is poor. As poor performance at the primacy position has been associated with hippocampal dysfunction (Baddeley & Warrington, 1970; Hermann et al., 1996; Milner, 1987) and hippocampal atrophy occurs early in the pathophysiology of AD (Convit et al. 1993), our results suggest that memory performance for the early list items can be a behavioral-based way to assess hippocampal integrity; and that poor delayed primacy recall performance should be considered as a sign of potential forthcoming generalized cognitive decline.

Funding

This study was funded in part by a grant to NP from Rockland County, NY, to support the Memory Evaluation Research Initiative (MERI). The authors have no conflict of interest to disclose.

Acknowledgments

DB is now at Liverpool Hope University, UK.

References

Albert M., Blacker, D., Moss, M. B., Tanzi, R., & McArdle, J. J. (2007). Longitudinal change in cognitive performance among individuals with mild cognitive impairment. *Neuropsychology, 21*, 158-169.

Albert, M., Moss, M. B., tanzi, R., & Jones, K. (2001). Preclinical prediction of AD using neuropsychological tests. *Journal of the International Neuropsychological Society, 7*, 631-639.

Azizian, A., & Polich, J. (2007). Evidence for attentional gradient in the serial position memory curve from event-related potentials. *Journal of Cognitive Neuroscience, 19*, 2071-2081.

Baddeley A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9*, 176-189.

Bayley, P. J., Salmon, D. P., Bondi, M. W., Bui, B. K., Olichnev, J., Delis, D. C., Thomas, R. G., Thal, L. J. (2000). Comparison of the serial position effect in very mild Alzheimer's disease, mild Alzheimer's disease, and amnesia associated with electroconvulsive therapy. *Journal of the International Neuropsychology Society, 6*, 290-298.

Bigler, E. D., Rosa, L., Schultz, F., Hall, S., & Harris, J. (1989). Rey-auditory verbal learning and Rey-osterrieth complex figure design performance in Alzheimer’s disease and closed head injury. *Journal of Clinical Psychology, 45*, 277-280.

Berti, V., Mosconi, L., Glodzik, L., Li, Y., Murray, J., De Santi, S., Pupi, A., Tsui, W., & DeLeon, M.J., (*in press*). Structural brain changes in normal individuals with a maternal history of Alzheimer’s. *Neurobiology of Aging*.

Blennow, K, DeLeon, M.J., & Zetterberg, H. (2006). Alzheimer’s disease. *Lancet. 368*, 387-403.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of

memory. *Psychological Review, 114*, 539-576.

Carlesimo GA, Fadda L, Sabbadini M, Caltagirone C. (1996). Recency effect in Alzheimer's disease: a reappraisal. *Quarterly Journal of Experimental Psychology, 49*, 315-325.

Convit, A., de Leon, M.J., Golomb, J., George, A.E., Tarshish, C.Y., Bobinski, M., Tsui, W., De Santi, S., Wegiel, J., & Wisniewski, H., 1993. Hippocampal atrophy in early Alzheimer's disease: anatomic specificity and validation. Psychiatric Quarterly, *64*, 371-87.

Farrer, L.A., Cupples, L.A., Haines, J.L., Hyman, B., Kukull, W.A., Mayeux, R., Myers, R.H., Pericak-Vance, M.A., Risch, N., & van Duijn, C.M. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *Journal of the American Medical Association, 278*, 1349-1356.

Foldi, N. S., Brickman, A. M., Schaefer, L. A., Kneutelska, M. E. (2003). Distinct serial position profiles and neuropsychological measures differentiate late-life depression from normal aging and Alzheimer's disease. *Psychiatry Research, 120*, 71-84.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Frederiksen, K.S., Garde, E., Skimminge, A., Ryberg, C., Rostrup, E., Baaré, W.F., et al. (In Press). Corpus Callosum Atrophy in Patients with Mild Alzheimer's Disease. *Neurodegenerative Disease*.

Gainotti G., Marra C., Villa G., Parlato V., Chiarotti F. (1998). Sensitivity and specificity of some neuropsychological markers of Alzheimer dementia. Alzheimer Disease & Associate Disorders, *12*, 152-62.

Gelman A., & Hill J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower (Ed.), *The psychology of learning and motivation*. New York: Academic Press. Pp. 129-153.

Gomar, J. J, Bobes-Bascaran, M. T., Conejero-Goldberg, C., Davies, P., & Goldberg, T. E. (2011). Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. *Archives of General Psychiatry, 68,* 961-969.

Hermann, B. P., Seidenberg, M., Wyler, A., Davies, K., Christeson, J., Moran, M., & Stroup, E. (1996). The effects of human hippocampal resection on the serial position curve. *Cortex, 32*, 323-334.

Howieson, D. B., Mattek, N., Seeyle, A. M., Dodge, H. H., Wasserman, D., Zitzelberger, T., & Jeffrey, K. (2011). *Journal of Clinical and Experimental Neuropsychology, 33*, 292-299.

Jack, C.R. Jr., Petersen, R.C., Xu, Y.C., Waring, S.C., O'Brien, P.C., Tangalos, E.G., Smith, G.E., Ivnik, R.J., Kokmen, E., 1997. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. Neurology, *49*, 786-794.

Jack, C.R., Petersen, R.C., Xu, Y.C., O'Brien, P. C., Smith, G.E., Ivnik, R. J., Boeve, B. F., Waring, S. C., Tangalos, E.G., Kokmen, E., 1999. Prediction of AD with MRI based hippocampal volume in mild cognitive impairment. *Neurology, 52*, 1397-1403.

Killiany, R.J., Gomez-Isla, T., Moss, M., Kikinis, R., Sandor, T., Jolesz, F., Tanzi, R., Jones, K., Hyman, B.T., & Albert, M.S. (2000). Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. *Annals of Neurology, 47*, 430–439.

Knopman, D.S., Mosley, T.H., Catellier, D.J., Coker, L.H. (2009).Fourteen-year longitudinal study of vascular risk factors, APOE genotype, and cognition: the ARIC MRI Study. *Alzheimers Dementia, 5,* 207-214.

La Rue, A., Hermann, B., Jones, J. J., Johnson, S., Ashtana, S., & Sager, M. A. (2008). Effect of parental family history of Alzheimer's disease on serial position profiles. *Alzheimer's and Dementia, 4,* 285-290.

Light, L. (1991). Memory and aging: Four hypotheses in search of data. *Annual Review of Psychology, 42,* 333-376.

McGaugh, J.L. (2000). Memory: a century of consolidation. *Science, 287*, 248-251.

Milner, B. (1987). Memory and the human brain. In M. Shafto (Ed.), *How we know*. San Francisco: Harper and Row. Pp. 31-59.

Mosconi, L., Berti, V., Swerdlow, R. H., Pupi, A., Duara, R., & de Leon, M. (2010). Maternal transmission of Alzheimer's disease: Prodromal metabolic phenotype and the search for genes. *Human Genomics, 3,* 170-193.

Murdock, B.B. (1960). The distinctiveness of stimuli. *Psychological Review, 67,* 16- 31.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64,* 482-488.

Neath, I. (2010). Evidence for similar principles in episodic and semantic memory: The presidential serial position function. *Memory & Cognition, 38,* 659-666.

Petersen, R.C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine, 256,* 183-194.

Prince, M., Cullen, M., & Mann, A. (1994). Risk factors for Alzheimer's disease and dementia: a case-control study based on the MRC elderly hypertension trial. Neurology, *44*, 97-104.

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89,* 63-77.

Rushby, J. A., Barry, R. J., & Johnstone, S. J. (2002). Event-related potential correlates of serial-position effects during and elaborative memory test. *International Journal of Psychophysiology, 46,* 13-27.

Salthouse, T.A. (1992). What do adult age differences in the Digit Symbol Substitution Test reflect? *Journal of Gerontology, 47,* 121-128.

Silverman, J.M., Ciresi, G., Smith, C.J., Marin, D.B., & Schnaider-Beeri, M. (2005). ariability of familial risk of Alzheimer disease across the late life span. *Archives of General Psychiatry, 62,* 565-573.

Tombaugh, T. N. (2005). Test-retest reliable coefficients and 5-year change scores for the MMSE and 3MS. *Archives of Clinical Neuropsychology, 20,* 485-503.

Verbeke, G., & Molenberghs, G. (2009). *Linear Mixed Models for Longitudinal Data*. G. Verbeke and G. Molenberghs (Eds.). New York: Springer.

Wilson, R.S., Barnes, L.L., Mendes de Leon, C.F., Aggarwal, N.T., Schneider, J.S., Bach, J., Pilat, J., Beckett, L.A., Arnold, S.E., Evans, D.A., & Bennett, D.A. (2002). Depressive symptoms, cognitive decline, and risk of AD in older persons. *Neurology, 59,* 364-370.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology, 55*, 235-269.

Yaffe, K., Weston, A., Graff-Radford, N. R., Satterfield, S., Simonsick, E. M., Younkin, S. G., Younkin, L. H., Kuller, L., Ayonayon, H. N., Ding, J., & Harris, T. B. (2011). Association of plasma beta-amyloid level and cognitive reserve with subsequent cognitive decline. *Journal of the American Medical Association, 305*, 261-266.

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neuroscience, 34*, 515-525.

Table. *Demographics. Number of subjects (i.e., N); Age in years (mean and standard deviation); MMSE score (with standard deviation); Gender (proportion of females); and Years of Education (with standard deviation), as a function of visit, when applicable.*

	Baseline	Visit 2	Visit 3	Visit 4	Visit 5
N	204	204	125	64	28
Age	71.5 (6.3)	73.3 (6.3)	75.0 (6.3)	76.4 (6.5)	77.3 (7.1)
MMSE	29.2 (0.7)	29.0 (1.7)	28.9 (2.2)	29.0 (2.5)	29.2 (1.4)
Gender	56%				
Years of Education	15.0 (3.1)				

Figure 1. Model-based estimates of expected MMSE scores at the 1, 3 and 5 year marks, assuming the average baseline MMSE score for the given number of primacy words. The three lines represent different numbers of primacy words recalled in the baseline delayed trial (4 words recalled, 2 words recalled, or 0 words recalled) with 95% confidence intervals. The dotted line represents the maximum MMSE score of 30.

Figure 2. Serial position curves for the delayed recall trial (a) and the learning trial (b). On the x-axis is reported the word serial position from 1 to 15. Vertical dotted lines separate the primacy region (words 1-4) from the middle region (words 5-11) and the recency region (words 12-15). On the y-axis, the proportion of correctly recalled words per position is reported. Two serial position curves are drawn; one curve reports the performance of subjects who either maintained the same level of MMSE over the subsequent visits, or improved (no decline), whereas the other curve depicts the performance of subjects who declined.