Title: Changes in visual memory in mild cognitive impairment. A longitudinal study with CANTAB.

Authors: María Campos-Magdaleno[1], David Leiva[2], Arturo X. Pereiro[1], Cristina Lojo-Seoane[1], Sabela C. Mallo[1], David Facal[1] & Onésimo Juncos-Rabadán[1]

Filiations: 1. Department of Developmental Psychology, University of Santiago de Compostela, Galicia, Spain. 2. Department of Methodology of Behavioural Sciences, University of Barcelona, Catalunya, Spain

Number of words: 5658

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain).
Email: maria.campos@usc.es

Abstract

Background: Mild cognitive impairment (MCI), as a stage in the cognitive continuum between normal aging and dementia, is mainly characterized by memory impairment. The aims of this study were to examine CANTAB measures of temporal changes of visual memory in MCI and to evaluate the usefulness of the baseline scores for predicting changes in cognitive status.

Methods: The study included 201 participants aged over 50 years with subjective cognitive complaints. Visual memory was assessed with four CANTAB tests (PAL, DMS, PRM and SSP) administered at baseline and on two further occasions, with a follow-up interval of 18-24 months. Participants were divided into three groups according to the change in their cognitive status: participants with subjective cognitive complaints who remained stable (SCC-Stable), MCI participants who remained stable (MCI-Stable) and MCI participants whose cognitive deterioration continued (MCI-Worsened). Linear Mixed Models were used to model longitudinal changes, with evaluation time as a fixed variable, and multinomial regression models were used to predict changes in cognitive status.

Results: Isolated significant effects were obtained for Age and Group with all CANTAB tests used. Interactions between Evaluation Time and Group were identified in the PAL and DMS tests, indicating different temporal patterns depending on the changes in cognitive status. Regression models also indicated that CANTAB scores were good predictors of changes in cognitive status.

Conclusions: Decline in visual memory measured by PAL and DMS tests can successfully distinguish different types of MCI, and considered together PAL, DMS, PRM and SSP can predict changes in cognitive status.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain).
Email: maria.campos@usc.es

Keywords: visual memory; CANTAB tests; mild cognitive impairment; linear mixed models; longitudinal assessment.

Introduction

Cognitive decline in the elderly can be considered a continuum ranging from a cognitively unimpaired state (CU), to the presence of subjective cognitive complaints (SCC) without objective cognitive impairment, also called Subjective Cognitive Decline (SCD) (Jessen et al., 2016; Molinuevo et al., 2017), followed by Mild Cognitive Impairment (MCI), characterized by presence of cognitive complaints, objective cognitive deterioration and preservation or minimal impairment of instrumental activities of daily living (Petersen, 2004; Petersen et al., 2018), and finally dementia, which is characterized by cognitive and behavioural symptoms that impair normal functioning in daily life (APA, 2013). The single and multiple domain subtypes of amnestic and non-amnestic MCI that involve deterioration in only one or in more than one cognitive domain may also represent different levels of cognitive decline, with the multiple domain subtype being the most extreme clinical state (Brambati et al., 2009; Han et al., 2012). Progression along the continuum is a complex process characterized by cognitive changes, transitions and diagnostic instability at SCD and MCI stages, conversion to dementia and recovery to CU (Facal, Guàrdia-Olmos & Juncos-Rabadán 2015; Petersen et al., 2018). However, taking the instability into account, MCI and the subtypes characterized by only memory impairments (amnestic single-domain) or by impairments in memory and in other cognitive domains (amnestic multi-domain) are considered high-risk states for progression to dementia, mainly Alzheimer's Disease (AD). Early detection of the different stages of cognitive decline and the progress of decline is a pressing research challenge in the prevention and treatment of dementia (Albert et al., 2011; Petersen et al., 2018).

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Previous studies have shown that visual memory impairment can differentiate MCI patients from cognitively unimpaired controls (Alescio-Lautier et al., 2007; Barbeau et al. 2008; Juncos-Rabadán, Facal, Pereiro & Lojo-Seoane, 2014a; Westerberg et al., 2013). Other studies have successfully predicted the progression from MCI to AD (De Anna et al., 2014; Defrancesco et al., 2013; Didic et al, 2013; Oltra-Cucarella et al., 2018; Reijs et al, 2017; Saxton et al., 2004) and even complete neurodegenerative progress from the cognitively impaired state to MCI and AD (Mistridis, Krumm, Monsch, Berres & Taylor, 2015). These findings indicate the importance of including reliable visual memory tests for diagnosing MCI and for studying the course of decline in different aspects of visual memory in progression to AD.

Computerized assessment of visual memory using the Cambridge Neuropsychological Test Automated Battery (CANTAB; Cambridge Cognition Ltd., 2012; Sahakian et al., 1988) has been used to differentiate controls, MCI and AD participants in cross-sectional studies (Alladi, Arnold, Mitchell, Nestor & Hodges 2006; de Rover et al., 2011; Juncos-Rabadán et al, 2014a; Junkkila, Oja, Laine & Karrasch, 2012; Swaison et al., 2001). CANTAB includes tests that assess visual Episodic Memory (EM) and visual Working Memory (WM). Both types of memory have been shown to be impaired early on in AD (Belleville, Sylvain-Roy, de Boysson & Ménard, 2008; Economou, Papageorgiou & Karageorgiou, 2011; van Geldrop et al., 2015). Deterioration in EM has been found to be a particularly strong predictor of progression to AD (Belleville et al., 2008; Landau et al., 2010).

Longitudinal evidence from research using the CANTAB visual memory tests remains scarce (Cacciamani et al., 2017; Juncos-Rabadán et al., 2016; Mitchell, Arnold, Dawson, Nestor & Hodges, 2009; Summers and Saunders, 2012). Summers and Saunders (2012) found that the decline in visual memory performance assessed with CANTAB measures (Paired Associates Learning, Spatial Span, Spatial Working Memory) in combination with the Rey

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Auditory Verbal Learning Test identified 100% of cases of MCI patients who progressed to AD after 20 months. However, Cacciamani et al. (2017) reported improvements in Spatial Working Memory, Spatial Recognition Memory and Paired Associated Learning after a follow-up period of 12 months in a small sample of MCI patients. Further investigation including larger sample sizes and longer intervals between assessments must be carried out to analyze the discriminant value and evolution of these memory measures.

The main purpose of the present study was to determine longitudinal patterns of performance of visual memory CANTAB tests in patients diagnosed at baseline with MCI and assessed twice with a follow-up interval of around 18 months to measure stability or deterioration of the condition. A secondary aim was to assess the usefulness of baseline CANTAB measures for predicting changes in cognitive status at the final follow-up stage.

Methodology

Participants

Participants were selected from the Compostela Aging Study (CompAS), an ongoing longitudinal project involving the detection and follow-up of Mild Cognitive Impairment in patients with subjective cognitive complaints and no prior diagnostic of dementia, psychiatric or neurological disorders attending primary care centres in Galicia, an autonomous region in northwest Spain (Juncos-Rabadán et al., 2012). We selected 201 patients aged over 50 years who had completed 3 visits (at Baseline, Time 1 and Time 2) with a between-test interval of around 18 months. The mean interval was 18.49 months (3.64 standard deviation, SD) between Baseline and Time 1, 17.72 months (3.81 SD) between Time 1 and Time 2, and 36.83 months (5.17 SD) between Baseline and Time 2. None of the participants had previously been diagnosed with MCI or dementia, clinical stroke, traumatic brain injury, motor-sensory defects, alcohol or drug abuse/dependence, or any neurological or psychiatric

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

disease. At baseline, participants were classified as single-domain amnestic MCI (sda-MCI), multiple-domain amnestic MCI (mda-MCI), single-domain non amnestic MCI (sdna-MCI) or multiple-domain non amnestic MCI (mdna-MCI), according to standard criteria (Albert et al., 2011; Dubois et al., 2007; Petersen, 2004). The criteria for diagnosis of MCI included the following: (a) self-reported, informant-corroborated concerns about cognition, assessed by a short version of the subjective memory complaints questionnaire (SMCQ; Benedet and Seisdedos, 1996); (b) performance of 1.5 standard deviations (SD) below age and education norms in one or more cognitive domains, assessed by the subscales of the Spanish version of the Cambridge cognitive examination, CAMCOG-R (Huppert et al., 1996; Spanish version: López-Pousa, 2003; Pereiro, Ramos-Lema, Juncos-Rabadán, Facal & Lojo-Seoane, 2015), except for memory, assessed by the short and long delay free recall from the Spanish version of the California verbal learning test (Delis et al., 1987; Spanish version: Benedet and Alejandre, 1998); (c) no significant or minimal impact on activities of daily living, assessed by instrumental activities of daily living scale (Lawton and Brody, 1969); and (d) the absence of dementia as established by the DSM-IV and NINCDS-ADRDA criteria. Participants performing as cognitively normal adults in general functioning and specific domain tests, according to norms by age and years of education, and presenting subjective cognitive complaints (SCC), were included in the SCC group. This group met the following criteria: (a) attending primary care health centres with self-reported cognitive concerns; and (b) confirmation of these concerns by the short Spanish version of the questionnaire for subjective memory complaints (Benedet and Seisdedos, 1996) administered to participants and a family member. The SCC group was considered a control group. All diagnoses were reached by consensus at a special meeting of the research team.

In each successive follow-up assessment, participants were reclassified as SCC, sda-MCI, mda-MCI, sdna-MCI, mdna-MCI and probable dementia (DSM-IV and NINCDS-

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

ADRDA) by applying the same criteria as at baseline. At the third evaluation, participants were classified into three groups according to the changes in their cognitive status: participants with SCC at Baseline who remained stable at Time 2 (SCC-stable group, n=148, 71.49%); participants diagnosed with MCI at Baseline who remained stable at Time 2 (MCI-stable group, n= 31, 15.45%); and participants diagnosed as sda-MCI or sdna-MCI at Baseline who progressed to mda-MCI, mdna-MCI or dementia at Time 1 or Time 2 (MCI-worsened group, n= 22, 13.04%). Probable AD or other types of dementia were diagnosed according to the DMS-IV and NINCDS-ADRDA criteria, and progression to dementia was confirmed by consultation of the medical history and recording the date of neurological diagnosis. We assumed, in accordance with Brambati et al. (2009) and Campos-Magdaleno, Díaz-Bóveda, Juncos-Rabadán, Facal & Pereiro (2016), that the change from single-domain to multiple-domain corresponds to cognitive worsening, in which multi-domain MCI represents the most severely impaired of the MCI subtypes.

All participants gave their written informed consent prior to participation in the study. The research project was approved by the Galician Ethics Committee for Clinical Research (Xunta de Galicia, Spain), and the study was performed in accordance with the ethical standards established in the 1964 Declaration of Helsinki and revised in Seoul 2008.

Materials and Procedure

Four CANTAB visual memory tests were administered: paired associates learning (PAL), pattern recognition memory (PRM), delayed matching to sample (DMS) and spatial span (SSP). The PAL test assesses visuospatial episodic memory and learning (Sahakian et al., 1988). One or more boxes containing a pattern are displayed on the screen and are opened in a random order. The patterns shown in the boxes are then displayed in the middle of the screen, one at a time, and participants are asked to touch the box in which the pattern was originally located. If the participant makes an error, the patterns are shown again as a reminder of the locations. The

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

level of difficulty (2, 4, 6 and 8 patterns) was increased throughout the tests. The outcome variable was the total number of errors adjusted to level 6, which represents a high level of difficulty and has been used by several researchers to study MCI and AD (Alladi et al., 2006; Chamberlain et al., 2011; Lenehan et al., 2016; Mitchell et al., 2009; Polcher et al., 2017). The PRM test assesses visual pattern recognition memory in a two-choice forced discrimination paradigm (Swainson et al., 2001). The participants were presented with two blocks of 12 visual patterns, each displayed separately. In the recognition phase, subjects are required to choose between a pattern they have already seen and a novel pattern. The outcome measure was the percentage of correct responses, considered in some previous studies as a specific episodic memory outcome (de Jager, Milwain & Budge, 2002; Juncos-Rabadán, Pereiro, Facal, Reboredo & Lojo-Seoane, 2014b; Nathan et al., 2017). Delayed matching to sample (DMS) assesses both simultaneous and short-term visual memory (Sahakian et al., 1988; Owen et al., 1993). Participants must select the pattern that exactly matches the sample from among four abstract choices that include distractors. In some trials, the sample and the choice patterns are shown simultaneously, while in others there is a delay of 0 ms, 4,000 ms or 12,000 ms. The outcome measure was the percentage of correct responses, also considered an episodic memory measure task (Juncos-Rabadán et al., 2014b, 2016; Sweeney, Kmiec & Kupfer, 2000). Spatial span (SSP) is a computerised version of the Corsi blocks task that assesses visual working memory capacity (Owen et al., 1990). A pattern of white squares is shown on the screen. Some of the squares change colour, one at a time, in a variable sequence. At the end of the presentation of each sequence, a tone indicates that the participant should touch each of the boxes in the same order that they were originally presented. The number of boxes in the sequence is increased from a level of two at the start of the test until a final level of nine, with three sequences at each level. The outcome variable, the span length, was calculated for the longest

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

sequence successfully recalled and was used as index for the SSP task (Saunders and Summers, 2010).

The four CANTAB tests were administered in a more extensive counterbalanced assessment carried out by trained psychologists. To control for the effect of visual acuity on performance of the CANTAB, we measured the visual acuity of both eyes with the Lighthouse near visual acuity test.

Statistical analysis

Cross-sectional analyses were carried out at baseline for socio-demographic and principal neuropsychological measures, which were modelled using non-parametric tests (e.g. Kruskal-Wallis and Mann-Whitney tests) to determine differences between groups, given the skewed empirical distributions and the small sample size in some cases. In order to model longitudinal changes in the CANTAB measures, we initially used (generalized) linear mixed models -(G)LMM- with random intercepts and random slopes. We considered that the intercepts might differ according to the memory trajectories of the participants and that different slopes would represent various temporal patterns of change in memory performance. We finally discarded random slopes in the estimated models due to convergence issues. The statistical models included the following independent variables or predictors as fixed effects: Evaluation Time (Baseline, Time 1 and Time 2), Group (SCC-stable, MCI-stable and MCI-worsened), and their interaction (Evaluation Time x Group). By specifying Group and Evaluation Time as fixed factors we can test pairwise comparisons of the estimated marginal means for the dependent variables for each group and at each evaluation time. As all models included random effects for intercepts and heteroskedasticity due to the group, the covariate age at baseline was standardized to enable interpretation of the intercept. Separate models were constructed for each dependent variable: PAL total errors adjusted for 6 shapes, PRM total percent correct, DMS total percent correct, and SSP span length. The SCC-stable group was considered the reference

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

group, and Baseline was considered the reference Evaluation Time. (G)LMMs assuming Gaussian response were used to model changes in percentages. (G)LMMs assuming Poisson response were used to model count data related to errors and SSP span length. When (G)LMM assumptions were not fulfilled (e.g. overdispersion of the data), a negative binomial distribution was used to model count data. A general procedure was used to model the relationship between responses and predictors: first, a null model including only the intercept was estimated (model 1); Group and Time predictors and their interaction were then gradually added in two subsequent models (2 and 3). Several goodness of fit indexes were used (e.g. Akaike's Information Criterion) to choose the best (G)LMMs for each response. In addition, we also modelled longitudinal changes in other cognitive outcomes, such as MMSE and CAMCOG-R scores, which clearly represent general cognitive performance, following the same procedures as with the CANTAB scores (see Supplementary Material S2 for more details).

GLMs were used to predict changes in cognitive status at the final follow-up stage by using the baseline CANTAB scores. Specifically, multinomial logistic regression models were used to assess the extent to which cognitive evolution groups at the final follow-up stage could be predicted by visual memory scores at baseline. Four multinomial logistic regression models were constructed with each of the CANTAB measures as predictors as well as a multiple regression model combining these measures as predictors. The age of participants was added as a covariate in all the abovementioned GLMs. Information criteria indices, such as AIC and BIC, were used to select the best candidate subset of predictors, as proposed by other authors (Fox, 2016; Weisberg, 2014); given that these indices are unbounded, the best fits are indicated by lower values. Thus, models with lowest AIC/BIC values were considered to provide the best fit to the data. The general criterion applied was selection of the model that showed, within the set of fit indicators, at least some positive evidence. For instance, a minimum difference in BIC of 2 units, which is equivalent to a minimum Bayes Factor of 3 (see Table 22.1 in Fox, 2016,

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

for further details), is considered supporting evidence for a specific model. As with the GLMs, AIC was used to assess the goodness of fit of the different models. Finally, the Area Under the Curve (AUC) index was estimated for all models in order to evaluate the predictive capacity of each.

Cross-sectional statistical analysis was performed with SPSS for Windows, version 21.0 (SPSS, Chicago, IL, USA). The (G)LMMs were constructed in R environment (version 3.6.2; R Core Team, 2019) with the nlme (version 3.1-143; Pinheiro, Bates, DebRoy, Sarkar & R Core Team, 2019) and lme4 packages (version 1.1-21; Bates et al., 2015).

Results

Socio-demographic and neuropsychological profiles of the groups at baseline are summarized in Table 1. Comparisons revealed no differences between groups in years of education and the Charlson Comorbidity Index (CCI). For the cognitive variables (except for the MMSE scores, which were similar in both MCI groups), the SCC-stable group performed best, followed by the MCI-stable group and MCI-worsened group. The MCI-worsened was the oldest group. Finally, MCI-stable had the highest scores in subjective cognitive complaints. No significant differences were found between groups in visual acuity. Results obtained with the (G)LMMs showed that cognitive decline was significantly more pronounced in the MCI groups (see Section S2 in Supplementary Material: (G)LMMs were estimated for MMSE and CAMCOG-R scores). Specifically, a significant interaction between Time and Group predictors was found in those models in which MMSE ($X^2(2)=21.50$; $p<.001$) and CAMCOG-R ($X^2(2)=19.99$; $p<.001$) scores were included as responses. The interaction can be summarized by the greater decrease in the general cognitive performance of the individuals included in MCI-worsened group than in the individuals included in the other two groups.

PAL total errors adjusted 6 shapes

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

We used (G)LMMs assuming a response according to a negative binomial distribution because of the presence of overdispersion (i.e. the spread parameter is significantly greater than the location parameter). Model 3, which included Evaluation Time, Group, Interaction Evaluation Time x Group, and the random effects for the intercepts, yielded the best fit (see Table 2). The results of Model 3 showed significant effects of the covariate Age [$X^2(1)=54.02$; $p<.001$], the variables Evaluation Time [$X^2(1)=14.72$; $p<.001$] and Group [$X^2(2)=75.81$; $p<.001$] and the Evaluation Time x Group interaction [$X^2(2)=108.83$; $p<.001$], indicating different temporal patterns in the two MCI and the SCC stable groups over time. Estimated means from the aforementioned model indicated that the scores of the SCC-Stable group scarcely changed over time (e. g. mean difference between Baseline and T2 = 1.5; p =.02) whereas the errors in the MCI-Stable and MCI-Worsened groups increased (Baseline-T2 means differences equal 12.07 and 36.99, respectively; $p<.001$). Figure 1 shows the estimated longitudinal trends for PAL total adjusted errors 6 shapes in the three groups across the three evaluation times.

INSERT HERE TABLE 2 AND FIGURE 1

PAL total errors adjusted 6 shapes at baseline also proved to be a good predictor of changes in cognitive status at the end of the follow-up ($X^2(2)=68.44$; $p<.001$). In this regard, the relative risk of being in the MCI-worsened group when PAL errors increased by one unit, relative to the reference SCC-stable group (see S1 section in Supplementary Material), was 1.035. The model including this variable as the only predictor displayed a good predictive capacity (AUC=0.78).

PRM total percent correct

GLMMs using normal response (Gaussian) for percentages showed that Model 2 (represented in Table 3) yielded a better fit than the other models. Model 2 included only random effects for the intercepts and fixed effect for Age at baseline ($X^2(1)=32.99$; $p<.001$),

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Evaluation Time ($X^2(1)=0.32$; p=.57) and Group ($X^2(2)=89.89$; p<.001). According to this model, Age at baseline and Group had significant effects, but the Time predictor did not have a significant effect. The latter predictor was retained in the model in order to estimate and show marginal means across time. Mean distributions indicated that the percentage of hits in PRM did not change over time, indicating that the initial differences between groups were maintained throughout evaluation times (SCC-stable> MCI-stable=MCI-worsened). Figure 1 represents the longitudinal trends for PRM total percent correct in the three groups across the three evaluation times.

INSERT HERE TABLE 3

PRM total percent correct at the baseline was found to be a useful predictor of changes in cognitive status at the end of the study period ($X^2(2)=75.49$; p<.001). Specifically, estimated multinomial logistic model (see section S1 in Supplementary Material) showed that by increasing the scores of this CANTAB test by one unit, the expected relative risk of being classified in the MCI-worsened group is 0.874 relative to the reference group, which was SCC-stable. The simple multinomial logistic model appeared to have a good predictive capacity (AUC=0.79).

DMS total percent correct

Model 3 yielded the best fit for percentages of correct responses in DMS obtained by means of GLMMs with Gaussian response (see Table 4), which included random effects for the intercepts and fixed effect for Age at baseline ($X^2(1)=66.63$; p<.001), Evaluation Time ($X^2(1)=0.32$; p=.57), Group ($X^2(1)=51.76$; p<.001) and the Time x Group interaction ($X^2(1)=22.96$; p<.001). According to this model, Age at baseline had a significant effect and, given the significant interaction, Group effect depends on Time and vice versa. In this regard, the distribution of the estimated means indicated a significant decline in the DMS percent

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

correct in the MCI-worsened group over time (Baseline-T2 means difference = 13.89; p<.001). By contrast, neither the SCC-stable group nor the MCI-stable group yielded significant differences when measurement times were compared (Baseline-T2 mean difference = -1.13; p=.38) (Baseline-T2 mean difference=-2.49; p=.43) (see Figure 1).

INSERT HERE TABLE 4

The multinomial logistic regression model using DMS total percent correct at baseline as the only predictor showed that this measure was useful for predicting the classification of individuals according to the change in cognitive status criteria ($X^2(2)=38.32$; p<.001). The relative risk ratio for being classified as MCI-worsened when the baseline DMS scores increased by one unit was 0.92 (see S1 section in Supplementary Material). The predictive capacity of the model can be regarded as good (AUC=0.71).

SSP span length

GLMMs using a Poisson response (i.e. the assumption of equidispersion was met) for SSP span length showed that Model 2 produced a better fit than the other alternatives. This model (see Table 5) included only random effects for the intercepts and fixed effect for Age at baseline ($X^2(1)=6.50$; p=.011), Evaluation Time ($X^2(1)=0.07$; p=.80) and Group ($X^2(2)=10.41$; p=.006). The Evaluation Time predictor was retained in the model in order to estimate and show marginal means across time. Considering the estimated marginal means of SSP span length on three measurement occasions (see Figure 1), significant differences were found between SCC-stable and MCI-worsened groups (mean differences in the three contrasts equal approximately 1.11; p <.01) but not between MCI-worsened and MCI-stable groups (three means differences close to -0.60; p>.05) or between the SCC-stable and MCI-stable groups (mean differences in the pairwise contrasts around 0.51; p>.05).

INSERT TABLE 5 HERE

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Inclusion of SSP span length in a multinomial logistic model to predict membership in the cognitive evolution groups led to observation of a significant effect (see section S1 of Supplementary Material; $X^2(2)=51.50$; $p<.001$). An increase of one unit in the baseline SSP span length score indicates that inclusion in the MCI-worsened group at the end of the study is less likely than being classified as SCC-stable (relative risk ratio is equal to approximately 0.12). The AUC (0.71) also indicates a good predictive capacity.

Combined CANTAB measures

Finally, we tested the predictive value of a set of predictors comprising the four CANTAB measures (i.e. PAL total errors adjusted 6 shapes, PRM total percent correct, DMS total percent correct and SSP span length) after controlling for age. The corresponding multinomial logistic model showed a significant effect of all CANTAB scores on the membership in cognitive evolution groups (see S1 section in Supplementary Material; Wald's tests for all estimated coefficients associated with CANTAB scores yielded $p < .05$) and the estimates were consistent with those included in the previous models including only one CANTAB score. The predictive capacity of the multinomial logistic model combining all CANTAB scores was very good (AUC=0.86).

Discussion

This study aimed to analyze the longitudinal patterns of performance of three visual episodic memory CANTAB tests and one visual working memory test in three diagnostic groups classified according to the changes in their cognitive status, and also to show the usefulness of the measures for predicting changes in the cognitive status of individuals at the end of the study. Overall, the results showed the existence of different patterns of longitudinal performance depending on the changes in diagnosis of the participants. Some CANTAB

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

outcomes differentiated participants who showed no cognitive impairment (SCC-Stable) and participants with MCI, and even between MCI participants who remained stable or worsened. The results indicate that assessing visual memory with CANTAB measures may be useful for differentiating between different stages of MCI in the cognitive continuum of dementia. Estimated simple and multiple multinomial logistic models used to assess the utility of CANTAB scores at the initial stage to predict cognitive evolution at the end of the study proved to have a good to very good predictive capacity (AUCs between 0.71 to 0.86; see section S1 of Supplementary Material for further information regarding the model estimates and performance). In summary, the models showed that the higher the visual memory score the lower the risk of being classified in the group with the worst cognitive outlook.

The age of participants at baseline significantly influenced performance of all tests over time. Older participants scored lower on all measures, regardless of the diagnostic group (SCC-stable, MCI-stable, MCI-worsened). The influence of age on the performance in the CANTAB visual memory tests of old adults with MCI and without cognitive impairment has been documented in cross sectional studies (Juncos-Rabadán et al., 2014a). The current findings add new evidence from a longitudinal design.

The study findings also show a main effect of Group, with the MCI-worsened group obtaining the worst scores in all CANTAB measures used at the three evaluation times. This group comprised participants with greater cognitive impairment, who were found to have progressed to multiple-domain MCI or dementia at either of the follow-up evaluations. The profile with worst performance in visual memory tests of multiple-domain MCI has already been shown in previous studies (Juncos-Rabadán et al., 2014b). Our results support the capacity of the CANTAB visual memory tests to show different performance profiles and discriminate between groups in the cognitive continuum from normal aging to dementia, and suggest the use of these tests for early diagnosis of cognitive impairment. The findings obtained with CANTAB

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

scores are consistent with some additional analyses done to verify that cognitive decline is significantly more pronounced in MCI groups. The findings showed that the changes differed significantly in the three study groups and that the individuals included in the MCI-worsened group showed the most negative changes in the general cognitive performance.

Regarding the main effect of the variable Evaluation Time, the PAL test was the only measure that indicated significant differences at the three evaluation moments in all participants. This significant main effect adds new evidence to previous studies on the utility of the PAL to assess visual memory and learning in old adults with and without cognitive impairment (Fowler, Saling, Conway, Semple & Louis, 2002; Junkkila et al., 2012; O'Connell et al., 2004; Polcher et al., 2017). Moreover, our results indicate that the PAL measure can detect changes in longitudinal performance related to evolution along a continuum of cognitive decline. Taking into account that longitudinal research is scarce, this finding is an important contribution and adds evidence to the pioneering work by Blackwell and colleagues (Blackwell et al., 2004), who observed that the same CANTAB measure was significantly correlated with the degree of subsequent cognitive deterioration in the early stages of AD.

The most interesting findings of the present study are the significant interactions between Evaluation Time x Group in the PAL and DMS measures. Regarding the PAL total errors adjusted-6 shapes, the interaction was significant for the MCI-stable and the MCI-worsened groups, indicating the existence of specific longitudinal patterns of performance for each. The marginal means indicate a small increase in errors in the SCC-stable group between the baseline and the follow-up evaluations, while in both MCI groups the errors increased significantly in the same periods. The increase was more important for the MCI-worsened group. The differences in PAL temporal patterns indicate a decline in the performance over time for all groups; however, they also enable discrimination between the least cognitively impaired group (SCC-stable) and the MCI groups, as well as between the MCI group that remain stable (MCI-

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

stable) and the MCI groups in which further deterioration occurs (MCI-worsened). Our findings add a new perspective to those reported by Cacciamani et al. (2017), who observed a marked improvement in PAL when comparing the baseline performance with the 6-month follow-up, but no difference in performance between 6- and 12-month follow-ups. This improvement may be the result of a practice effect due to the short follow-up period; however, the practice effect may disappear when longer follow-up intervals between PAL tests are used in longitudinal assessments.

Regarding the DMS, the Evaluation Time x Group interaction was only significant in the MCI-worsened group, in which the test performance declined over time. The performance of the other two groups, SSC-stable and MCI-stable, did not vary significantly. The Evaluation Time x Group interaction was not significant for either the PRM total percent correct or SSP span length. However, the estimated marginal means showed significant differences between SCC-stable and MCI-worsened groups, indicating a clear decline in the latter group over time.

The measures in which a significant Evaluation Time x Group interaction was observed correspond to the two CANTAB tests (PAL and DMS) most closely related to episodic memory (de Jager et al., 2002; Juncos-Rabadán et al., 2014a, 2014b, 2016; Nathan et al., 2017; Sweeney et al., 2000). PAL involves visuospatial episodic memory and learning, and DMS involves short-term memory of complex visual patterns. Decline in episodic memory has been described as one of the most potent predictors of progression to Alzheimer's disease (Belleville et al., 2008; Landau et al., 2010), and our results show that the PAL total errors adjusted-6 shapes and the DMS total percent correct enable detection of longitudinal changes that may be indicative of progression in the continuum of cognitive deterioration.

However, the measures the PRM total percent correct and the SSP span length that differed significantly between groups (Group main effect) did not indicate differences between groups over time (Evaluation Time x Group interaction). PRM involves memory and subsequent

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

recognition of sequences of visual patterns, which may be related to the attentional span capacity, which is associated with working memory. In previous studies, contradictory findings regarding span length as a measure of working memory that differentiates participants according to diagnosis and progression have been reported. While a large number of studies support the existence of impairment in span length prior to diagnosis of dementia (Belleville et al., 2017; Economou et al., 2006; Gagnon and Belleville, 2011; van Geldrop et al., 2015; Saunders and Summers, 2010), other studies obtained contradictory or non-meaningful results (Griffith et al., 2006; Guarch, Marcos, Salamero, Gastó, & Blesa, 2008; Kessels, Overbeek, & Bouman 2015), questioning the value of the measure for early detection of cognitive impairment. Our findings indicate that the PRM measure and the SSP span cannot differentiate longitudinal patterns between groups.

We conclude that visual episodic memory declines in people with MCI over time and that this decline may be a cognitive indicator of the progression in the continuum ranging from the stage characterized by presence of cognitive complaints without objective cognitive impairment to dementia, through the different levels of severity of MCI. PAL total errors adjusted-6 shapes outcome, and DMS percent correct total measures differentiate the changes in participants in the continuum of cognitive deterioration: people with and without objective deterioration, and people who worsen or remain stable over time. In addition, the between-evaluation intervals used in longitudinal studies should be wide enough to prevent practice effects.

Membership of groups characterized by change in cognitive status developed at the second follow-up stage (T 2) has proven to be accurate in the light of different types of evidence. First, a different pattern of change was observed in CANTAB measurements according to this classification. Secondly, different patterns of change were also observed in other cognitive scores such as MMSE and CAMCOG-R when comparing the groups included in this study. Finally, comparison of membership in groups obtained by the procedure described in this study

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

with a classification obtained by means of non-parametric clustering of multivariate trajectories (i.e. individual trajectories in the 4 CANTAB scores) revealed a similarity index of 0.74, which indicates a good level of agreement. In summary, we demonstrated that the visual CANTAB scores a) are useful for predicting cognitive evolution in the time-period included in this study, b) differ over time depending on the change in cognitive status of individuals, and c) allow researchers to classify individuals consistently in comparison with other cognitive outcomes (i.e. clinical assessment at the second follow-up).

The limitations of the present study include the fact that only one group of patients with MCI that worsened over time was considered. By not having a larger number of participants in whom deterioration tended to worsen, it was not possible to differentiate people who progress to multiple-domain MCI from those who progress to dementia, and both were included within the same group. This hinders interpretation of the results, as although the participants progress in the same direction of the continuum of cognitive deterioration, they show important differences regarding the degree of cognitive impairment and functional capacity. Differences between both types of participants in their CANTAB longitudinal profiles should be considered in future studies. On the other hand, the interval of thirty-six months between baseline and the final evaluation may not be long enough for full assessment of the progress. We hope in the future to be able to collect longitudinal data over a longer period of time, as the current longitudinal research is still ongoing. We expect to conduct a third follow-up evaluation to assess changes that have occurred in a period of approximately 54 months (4.5 years) after baseline.

Financial support

This research was supported through FEDER founds by the Spanish Directorate General of Scientific and Technical Research (Project Ref. PSI2014-55316-C3-1-R), the National Research Agency (Spanish Ministry of Science, Innovation and Universities) (Project Ref.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

PSI2017-89389-C2-1-R) and by the Galician Government (Consellería de Cultura, Educación e Ordenación Universitaria; axudas para a consolidación e estruturación de unidades de investigación competitivas do Sistema Universitario de Galicia; GI-1807-USC: Ref. ED431-2017/27).

Conflicts of interest.

None.

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in Seoul 2008.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., … Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association work groups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia: The Journal of the Alzheimer's Association, 7*(3), 270-279. Doi: 10.1016/j.jalz.2011.03.008

Alescio-Lautier, B., Michel, B. F., Herrera, C., Elahmadi, A., Chambon, C., Touzet, C., & Paban, V. (2007). Visual and visuospatial short-term memory in mild cognitive impairment and Alzheimer disease: role of attention. *Neuropsychologia, 45*(8), 1948-1960. Doi: 10.1016/j.neuropsychologia.2006.04.033

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Alladi, S., Arnold, R., Mitchell, J., Nestor, P. J., & Hodges, J. R. (2006). Mild cognitive impairment: applicability of research criteria in a memory clinic and characterization of cognitive profile. *Psychological Medicine, 36*(4), 507-515. Doi: 10.1017/S0033291705006744

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. American Psychiatric Association: Washington, DC.

Barbeau, E. J., Ranjeva, J. P., Didic, M., Confort-Gouny, S., Felician, O., Soulier, E., … Poncet, M. (2008). Profile of memory impairment and gray matter loss in amnestic mild cognitive impairment. *Neuropsychologia, 46*(4), 1009-1019. Doi: 10.1016/j.neuropsychologia.2007.11.019

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models usign lme4. *Journal of Statistical Software*, *67*(1), 1-48. Doi:10.18637/jss.v067.i01

Belleville, S., Fouquet, C., Hudon, C., Zomahoun, H. T. V., Croteau, J., & Consortium for the Early Identification of Alzheimer's disease-Quevec. (2017). Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: a systematic review and meta-analysis. *Neuropsychology Review,* 27(4), 328-353. Doi: 10.1007/s11065-017-9361-5

Belleville, S., Sylvain-Roy, S., de Boysson, C., & Ménard, M. C. (2008). Characterizing the memory changes in persons with mild cognitive impairment. *Progress in Brain Research, 169*, 365-375. Doi: 10.1016/S0079-6123(07)00023-4

Benedet, M. J., & Alejandre, M. A. (1998). *TAVEC: Test de Aprendizaje Verbal de España-Complutense*. Madrid: TEA ediciones.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Benedet, M. J., & Seisdedos, N. (1996). *Evaluación Clínica de las Quejas de Memoria en la Vida Cotidiana.* Madrid: Editorial Médica Panamericana.

Blackwell, A. D., Sahakian, B. J., Vesey, R., Semple, J. M., Robbins, T. W., & Hodges, J. R. (2004). Detecting dementia: novel neuropsychological markers of preclinical Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 17*(1-2), 42-48. Doi: 10.1159/000074081

Brambati, S. M., Belleville, S., Kergoat, M. J., Chayer, C., Gauthier, S., Joubert, S. (2009). Single- and multiple-domain amnestic mild cognitive impairment: two sides of the same coin? *Dementia and Geriatric Cognitive Disorders, 28*(6), 541–549. Doi: 10.1159/000255240

Cacciamani, F., Salvadori, N., Eusebi, P., Lisetti, V., Luchetti, E., Calabresi, P., & Parnetti, L. (2018). Evidence of practice effect in CANTAB spatial working memory test in a cohort of patients with mild cognitive impairment. *Applied Neuropsychology: Adult, 25*(3), 237-248. Doi:10.1080/23279095.2017.1286346

Campos-Magdaleno, M., Díaz-Bóveda, R., Juncos-Rabadán, O., Facal, D., & Pereiro, A. X. (2016). Learning and serial effects on verbal memory in mild cognitive impairment. *Applied Neuropsychology: Adult, 23*(4), 237-250. Doi: 10.1080/23279095.2015.1053887

Chamberlain, S. R., Blackwell, A. D., Nathan, P. J., Hammond, G., Robbins, T. W., Hodges, J. R., … Sahakian, B. J. (2011). Differential cognitive deterioration in dementia: a two year longitudinal study. *Journal of Alzheimer's Disease, 24*(1), 125-136. Doi: 10.3233/JAD-2010-100450.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

De Anna, F., Felician, O., Barbeau, E., Mancini, J., Didic, M., & Ceccaldi, M. (2014). Cognitive changes in mild cognitive impairment patients with impaired visual recognition memory. *Neuropsychology, 28*(1), 98–105. Doi: 10.1037/neu0000032

De Jager, C. A., Milwain, E., & Budge, M. (2002). Early detection of isolated memory deficits in the elderly: the need for more sensitive neuropsychological tests. *Psychological Medicine, 32*(3), 483-491. Doi: 10.1017/S003329170200524X

De Rover, M., Pironti, V. A., McCabe, J. A., Acosta-Carbonero, J., Arana, F. S., Morein-Zamir, S., et al. (2011). Hippocampal dysfunction in patients with mild cognitive impairment: a functional neuroimaging study of a visuospatial paired associates learning task. *Neuropsychologia, 49*(7), 2060-2070. Doi: 10.1016/j.neuropsychologia.2011.03.037.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. (1987). *California Verbal Learning Test*. San Antonio, TX: Psychological Corporation.

Defrancesco, M., Marksteiner, J., Deisenhammer, E., Kemmler, G., Djurdjevic, T., & Schocke, M. (2013). Impact of white matter lesions and cognitive deficits on conversion from mild cognitive impairment to Alzheimer's disease. *Journal of Alzheimer's Disease, 34*(3), 665-672. Doi: 10.3233/JAD-122095

Didic, M., Felician, O., Barbeau, E. J., Mancini, J., Latger-Florence, C., Tramoni, E., & Ceccaldi, M. (2013). Impaired visual recognition memory predicts Alzheimer's disease in amnestic mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders, 35*(5-6), 291-299. Doi: 10.1159/000347203

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., … Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurology, 6*(8), 734-746. Doi: 10.1016/S1474-4422(07)70178-3

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Economou, A., Papageorgiou, S., & Karageorgiou, C. (2006). Working-delayed memory difference detects mild cognitive impairment without being affected by age and education. *Journal of Clinical and Experimental Neuropsychology, 28*(4), 528-535. Doi: 10.1080/13803390590949340

Facal, D., Guàrdia-Olmos, J., & Juncos-Rabadán, O. (2015). Diagnostic transitions in mild cognitive impairment by use of simple Markov models. *International Journal of Geriatric Psychiatry, 30*(7), 669-676. doi: 10.1002/gps.4197

Fowler, K, S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. (2002). Paired associate performance in the early detection of DAT. *Journal of the International Neuropsychological Society, 8*(1), 58-71. Doi: 10.1017/S1355617702811067

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models* (3rd Edition). Thousand Oaks, CA: SAGE.

Gagnon, L. G., & Belleville, S. (2011). Working memory in mild cognitive impairment and Alzheimer's Disease: contribution of forgetting and predictive value of complez span tasks. *Neuropsychology, 25*(2), 226-236. Doi: 10.1037/a0020919.

Griffith, R. H., Netson, K. L., Harrell, L. E., Zamrini, E. Y., Brockington, J. C., Marson, D. C. (2006). Amnestic mild cognitive impairment: diagnostic outcomes and clinical prediction over a two-year time period. *Journal of the International Neuropsychological Society, 12*(2), 166-175. Doi: 10.1017/S1355617706060267

Guarch, J., Marcos, T., Salamero, M., Gastó, C., & Blesa, R. (2008). Mild cognitive impairment: a risk indicator of later dementia, or a preclinical phase of the disease? *International Journal of Geriatric Psychiatry, 23*(3), 257-265. Doi: 10.1002/gps.1871

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Han, J. W., Kim, T. H., Lee, S. B., Park, J. H., Lee, J. J., … Kim K. W. (2012). Predictive validity and diagnostic stability of mild cognitive impairment subtypes. *Alzheimer's & Dementia. 8*, 553–559. Doi: 10.1016/j.jalz.2011.08.007

Huppert, F., Jorm, A. F., Brayne, C., Girling, D. M., Barkely, C., Beardsall, L., & Paykel, E. S. (1996). Psychometric properties of the CAMCOG. *Ageing, Neuropsychology, and Cognition, 3,* 1-4. Doi: 10.1080/13825589608256624

Jessen, F., Amariglio, R. E., van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., … Subjective Cognitive Decline Initiative (SCD-I) Working Group. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's and Dementia, 10*(6), 844-852. Doi: 10.1016/j.jalz.2014.01.001

Juncos-Rabadán, O., Pereiro, A. X., Facal, D., Rodríguez, N., Lojo, C., Caamaño, J. A., ... Eiroa, P. (2012). Prevalence and correlates of cognitive impairment in adults with subjective cognitive complaints in primary care centres. *Dementia and Geriatric Cognitive Disorders, 33*(4), 226-232. Doi: 10.1159/000338607

Juncos-Rabadán, O., Facal, D., Pereiro, A. X., & Lojo-Seoane, C. (2014a). Visual memory profiling with CANTAB in mild cognitive impairment (MCI) subtypes. *International Journal of Geriatric Psychiatry, 29*(10), 1040-1049. Doi: 10.1002/gps.4095

Juncos-Rabadán, O., Pereiro, A. X., Facal, D., Reboredo, A., & Lojo-Seoane, C. (2014b). Do the Cambridge Neuropsychological Test Automated Battery episodic memory measures discriminate amnestic mild cognitive impairment? *International Journal of Geriatric Psychiatry, 29*(6), 602-609. Doi: 10.1002/gps.4042

Juncos-Rabadan, O., Pereiro, A. X., Facal, D., Lojo-Seoane, C., Mallo, S. C., & Campos-Magdaleno, M. (20 16). Longitudinal changes in visual memory in mild cognitive

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

impairment versus normal aging in people with subjective cognitive complaint. *Alzheimer's & Dementia, 12*(7), P754-P755. Doi: 10.1016/j.jalz.2016.06.1438

Junkkila, J., Oja, S., Laine, M., Karrasch, M. (2012). Applicability of the CANTAB-PAL computerized memory test in identifying amnestic mild cognitive impairment and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 34*(2), 83-89. Doi: 10.1159/000342116

Kessels, R. P. C., Overbeek, A., & Bouman, Z. (2015). Assessment of verbal and visuospatial working memory in mild cognitive impairment and Alzheimer's dementia. *Dementia & Neuropsychologia, 9*(3), 301-305. Doi: 10.1590/1980-57642015DN93000014

Landau, S. M., Harvey, D., Madison, C. M., Reiman, E. M., Foster, N. I., Aisen, P. S., … Jagust, W. J. (2010). Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology, 75*(3), 230-238. Doi: 10.1212/WNL.0b013e3181e8e8b8

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist, 9*(3), 179-186.

Lenehan, M. E., Summers, M. J., Saunders, N. L., Summers, J. J., & Vickers, J. C. (2016). Does the Cambridge Automated Neuropsychological Test Battery (CANTAB) distinguish between cognitive domains in healthy older adults?. *Assessment*, 23(2), 163-172. Doi: 10.1177/1073191115581474.

Mistridis, P., Krumm, S., Monsch, A. U., Berres, M., & Taylor, K. I. (2015). The 12 years preceding mild cognitive impairment due to Alzheimer's disease: the temporal emergence of cognitive decline. *Journal of Alzheimer's Disease, 48*(4), 1095-1107. Doi: 10.3233/JAD-150137.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Mitchell, J., Arnold, R., Dawson, K., Nestor, P. J., & Hodges, J. R. (2009). Outcome in subgrupos of mild cognitive impairment (MCI) is highly predicatble using a sample algorithm. *Journal of Neurology, 259*(9), 1500-1509. Doi: 10.1007/s00415-009-5152-0

Molinuevo, J. L., Rabin, L. A., Amariglio, R., Buckley, R., Dubois, B., Ellis, K. A. (2017). Implementation of subjective cognitive decline criteria in research studies. *Alzheimer's & Dementia, 13*(3), 296-311. Doi: 10.1016/j.jalz.2016.09.012

Nathan, P. J., Lim, Y. Y., Abbott, R., Galluzzi, S., M arizzoni, M., Babiloni, C., … PharmaCog Consortium. (2017). Association between CSF biomarkers, hippocampal volume and cognitive function in patients with amnestic mild cognitive impairment (MCI). *Neurology of Aging, 53*, 1-10. Doi: 10.1016/j.neurobiolaging.2017.01.013.

O'Connell, H., Coen, R., Kidd, N., Warsi, M., Chin, A. V., & Lawlor, B. A. (2004). Early detection of Alzheimer's disease (AD) using the CANTAB paired Associates Learning Test. *International Journal of Geriatric Psychiatry, 19*(12), 1207-1208. Doi: 10.1002/gps.1180

Oltra-Cucarella, J., Sánchez-Sansegundo, M., Lipnicki, D. M., Crawford, J. D., Lipton, R. B., Katz, M. J., … Cohort Studies of Memory in an International Consortium (COSMIC). (2018). Visual memory tests enhance the identification of amnestic MCI cases at greater risk of Alzheimer's disease. *International Psychogeriatrics, 25*, 1-10. Doi: 10.1017/S104161021800145X.

Owen, A. M., Beksinska, M., Jamnes, M., Leigh, P. N., Summers, B. A., Marsden, C. D., … Robbins, T. W. (1993). Visuospatial memory deficits at different stages of Parkinson's disease. *Neuropsychologia, 31*(7), 627-644. Doi: 10.1016/0028-3932(93)90135-M

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Pereiro, A. X., Ramos-Lema, S., Juncos-Rabadán, O., Facal, D., & Lojo-Seoane, C. (2015). Normative scores of the Cambridge Cognitive Examination-Revised in healthy Spanish population. *Psicothema, 27*(1), 32-39. Doi: 10.7334/psicothema2014.169

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

Petersen, R. C., Lopez, O., Armstrong, M. J., Getchius, T. S., Ganguli, M., Gloss, D., ... Sager, M. (2018). Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. *Neurology, 90*(3), 126-135. Doi: 10.1212/WNL.0000000000004826

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team (2018). nlme: Linear and nonlinear mixed effects models [Computer Software]. Retrieved from https://CRAN.R-project.org/package=nlme.

Polcher, A., Frommann, I., Koppara, A., Wolfsgruber, S., Jessen, F., Wagner, M. (2017). Face-name associative recognition deficits in subjective cognitive decline and mild cognitive impairment. *Journal of Alzheimer's Disease, 56*(3), 1185-1196. Doi: 10.3233/JAD-160637

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria [Computer Software]. Retrieved from https://www.R-project.org/.

Reijs, B. L., Ramakers, I. H., Köhler, S., Teunissen, C. E., Koel-Simmelink, M., Nathan, P. J., ... Vandenberghe, R. (2017). Memory correlates of Alzheimer's disease cerebrospinal fluid markers: a longitudinal cohort study. *Journal of Alzheimer's Disease, 60*(3), 1119-1128. Doi: 10.3233/JAD-160766.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Sahakian, B. J., Morris, R. G., Evenden, J. L., Heald, A., Levy, R., Philpot, M., & Robbins, T. W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain, 111*, 695-718. Doi: 10.1093/brain/111.3.695

Saunders, N. L. J., & Summers, M. J. (2010). Attention and working memory deficits in mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology, 32*(4), 350-357. Doi: 10.1080/13803390903042379

Saxton, J., Lopez, O. L., Ratcliff, G., Dulberg, C., Fried, L. P., Carlson, M. C., ... Kuller, L. (2004). Preclinical Alzheimer disease: neuropsychological test performance 1.5 to 8 years prior to onset. *Neurology, 63*(12), 2341-2347. Doi: 10.1212/01.WNL.0000147470.58328.50

Summers, M. J., & Saunders, N. L. J. (2012). Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. *Neurpsychology, 26(*4), 498-508. Doi: 10.1037/a0028576

Swaison, R., Hodges, J. R., Galton, C. J., Semple, J., Michael, A., Dunn, B. D., … Sahakian, B. J. (2001). Early detection and differential diagnosis of Alzheimer's Disease and depression with neuropsychological tasks. *Dementia and Geriatric Cognitive Disorders, 12*(4), 265-280. Doi: 10.1159/000051269

Sweeney, J. A., Kmiec, J. A., & Kupfer, D. J. (2000). Neuropsychologic impairments in bipolar and unipolar mood disorders on the CANTAB neurocognitive battery. *Biological Psychiatry, 48*(7), 674-684. Doi: 10.1016/S0006-3223(00)00910-0

Van Geldrop, B., Heringa, S. M., van den Berg, E., Olde Rikkert, M. G., Biessels, G. J., & Kessels, R. P. (2015). Working memory binding and episodic memory formation in aging, mild cognitive impairment, and Alzheimer's dementia. *Journal of Clinical and Experimental Neuropsychology, 37*(5), 538-48. Doi: 10.1080/13803395.2015.1037722.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Weisberg, S. (2014). *Applied Linear Regression* (4th Edition). Hoboken, NY: Wiley

Westerberg, C., Mayes, A., Florczak, S. M., Chen, Y., Creery, J., Parrish, T., Weintraub, S., … Paller, K. A. (2013). Distinct medial temporal contributions to different forms of recognition in amnestic mild cognitive impairment and Alzheimer's disease. *Neuropsychologia, 51*(12), 2450-2461. Doi: 10.1016/j.neuropsychologia.2013.06.025

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Table 1. Mean and standard deviations (in parentheses) of the demographic and neuropsychological measures at baseline for the three groups: subjective cognitive complaints (SCC) that remain stable (SCC-stable); mild cognitive impairment that remain stable (MCI-stable); mild cognitive impairment that worsened (MCI-worsened).

	SCC-stable Group 1 N=149	MCI-stable Group 2 N=32	MCI-worsened Group 3 N=27	Kruskal Wallis χ^2(gl)	Group comparison
Age	64.26 (8.83) Range:50-87	70.94 (7.54) Range: 54-83	75.44 (7.14) Range: 61-87	39.46 (2)**	G3 > G2 > G1
Gender	Women: 70.3% Men: 29.7%	Women: 68.8% Men: 31.3%	Women: 55.6% Men: 44.4%		
Years of Education	10.28 (4.71) Range: 2-22	9.15 (3.40) Range: 2-17	9.30 (4.79) Range: 4-25	1.09 (2)	
SCC	18.84 (4.54) Range: 7-31	20.25 (4.09) Range: 10-32	18.07 (4.64) Range: 13-33	6.83 (2)*	G2 > G1, G3
Lawton-Brody	7.55 (.95) Range: 4-8	6.8 (1.55) Range: 3-8	6.15 (2.08) Range: 2-8	17.29 (2)**	G1 > G2, G3
CCI	.76 (.84) Range: 0-3	1.09 (1.02) Range: 0-4	.70 (.86) Range:0-3	3.65 (2)	
MMSE	28.34 (1.34)	25.13 (2.89)	24.04 (2.53)	73.39 (2)**	G1 > G2, G3
CAMCOG	89.88 (6.96)	77.40 (8.99)	70.92 (10.08)	81.74 (2)**	G1 > G2 > G3
CVLT-SDFR	11.01 (2.50)	4.50 (3.00)	2.37 (2.04)	113.20 (2)**	G1 > G2 > G3
CVLT-LDFR	11.85 (2.57)	5.53 (3.77)	2.48 (2.43)	105.13 (2)**	G1 > G2 > G3
Visual Acuity	.55 (.17) Range: .20-1.00	.52 (.16) Range: .20-.80	.53 (.18) Range: .33-.80		

Note: MMSE= MiniMental State Examination CCI=Charlson Comorbidity Index. SCC: Subjective Cognitive Complaints (patient). CAMCOG= Cambridge Cognitive Examination (total score). CVLT SDFR= California Verbal Learning Test, Short Delay Free Recall. CVLT LDFR= California Verbal Learning Test, Long Delay Free Recall. Visual Acuity = Lighthouse test.

*= $p < .05$

**= $p < .01$

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Table 2. Summary of models compared for PAL total errors adjusted-6 shapes. All models include random effects for intercepts and age at baseline as covariate. Model 1 is the null mixed model (i.e. random intercepts and age covariate only); Model 2 is the mixed model with main effects; Model 3 is the mixed model with main effects and interactions. Coefficients and standard errors (in parentheses) are shown on a log scale of number of errors (i.e. natural log of the response).

	Dependent variable: PAL total errors adjusted-6 shapes		
	Model 1	Model 2	Model 3
Age at baseline	0.559*** (0.054)	0.379*** (0.051)	0.377*** (0.051)
Evaluation Time		0.036*** (0.009)	-0.032*** (0.012)
MCI-worsened		1.065*** (0.153)	0.888*** (0.156)
MCI-stable		0.904*** (0.131)	0.780*** (0.134)
Evaluation Time x MCI-worsened			0.280*** (0.030)
Evaluation Time x MCI-stable			0.137*** (0.021)
Intercept	3.403*** (0.054)	3.105*** (0.056)	3.170*** (0.057)
Observations	624	624	624
Log Likelihood	-3,005.318	-2,965.740	-2,911.160
Akaike Inf. Crit.	6,016.637	5,943.481	5,838.320
Bayesian Inf. Crit.	6,029.561	5,969.329	5,872.785

Note: ***$p < 0.01$

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Table 3. Summary of model comparison for PRM total percent correct. All models include random effects for intercepts and age at baseline as covariate. Model 1 is the null mixed model (i.e. random intercepts and age covariate only); Model 2 is the mixed model with main effects; Model 3 is the mixed model with main effects and interactions. Coefficients and standard errors (in parentheses)

	Dependent variable: PRM percent correct total		
	Model 1	Model 2	Model 3
Age at baseline	-5.430*** (0.656)	-3.439*** (0.599)	-3.460*** (0.599)
Evaluation Time		0.213 (0.379)	0.405 (0.405)
MCI-worsened		-16.591*** (2.362)	-16.673*** (2.823)
MCI-stable		-12.344*** (1.645)	-10.625*** (1.967)
Evaluation Time x MCI-worsened			0.286 (2.471)
Evaluation Time x MCI-stable			-1.974 (1.249)
Intercept	83.072*** (0.649)	85.735*** (0.731)	85.546*** (0.744)
Observations	560	560	560
Log Likelihood	-2,055.846	-2,014.233	-2,010.012
Akaike Inf. Crit.	4,123.691	4,046.466	4,042.024
Bayesian Inf. Crit.	4,149.637	4,085.337	4,089.493

Note: ***$p < 0.01$

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Table 4. Summary of compared models for DMS total percent correct. All models include random effects for intercepts and age at baseline as covariate. Model 1 is the null mixed model (i.e. random intercepts and age covariate only); Model 2 is the mixed model with main effects; Model 3 is the mixed model with main effects and interactions. Coefficients and standard errors (in parentheses)

	Dependent variable: DMS total percent correct		
	Model 1	Model 2	Model 3
Age at baseline	-6.170*** (0.565)	-4.555*** (0.554)	-4.456*** (0.546)
Evaluation Time		0.212 (0.387)	0.563 (0.423)
MCI-worsened		-12.014*** (1.831)	-7.677*** (2.038)
MCI-stable		-5.990*** (1.470)	-6.597*** (1.738)
Evaluation Time x MCI-worsened			-7.506*** (1.610)
Evaluation Time x MCI-stable			-0.684 (1.096)
Intercept	78.252*** (0.565)	80.170*** (0.707)	79.847*** (0.720)
Observations	555	555	555
Log Likelihood	-2,001.541	-1,975.448	-1,961.768
Akaike Inf. Crit.	4,011.083	3,964.896	3,941.536
Bayesian Inf. Crit.	4,028.344	3,995.065	3,980.292

Note: ***$p < 0.01$

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

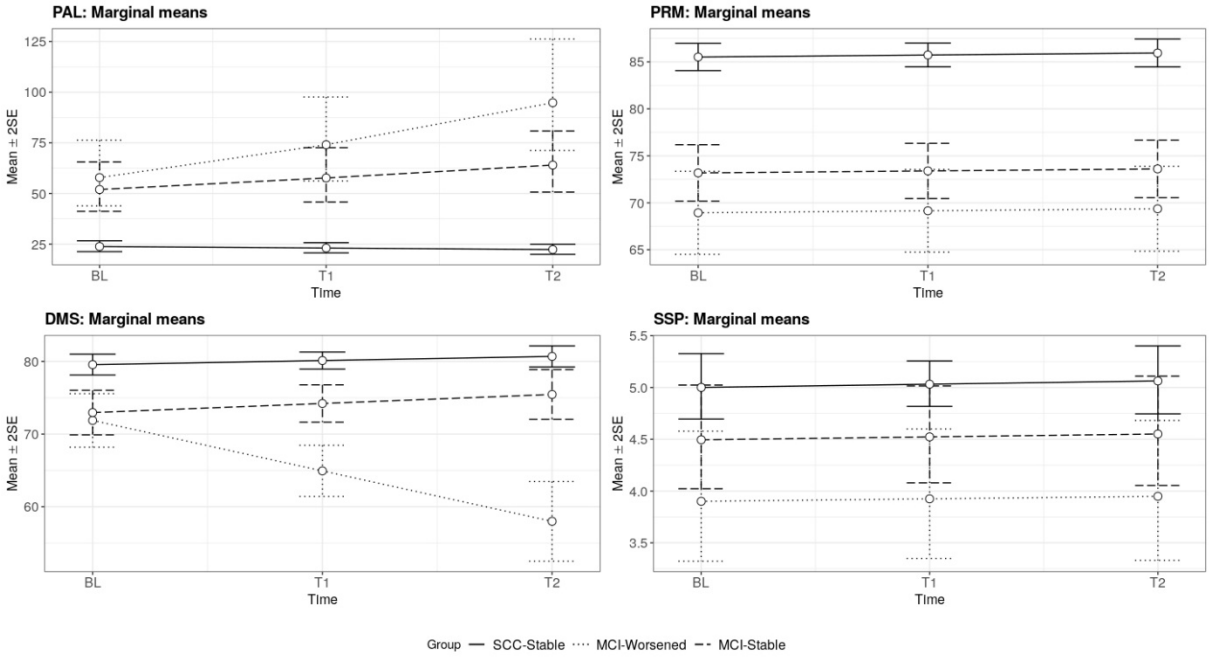
Table 5. Summary of models compared for SSP span length. All models include random effects for intercepts and age at baseline as covariate. Model 1 is the null mixed model (random intercepts and age covariate only); Model 2 is the mixed model with main effects; Model 3 is the mixed model with main effects and interactions. Coefficients and standard errors (in parentheses) are shown on the log scale of number of correct responses (i.e. natural log of the response).

	Dependent variable: SSP span length		
	Model 1	Model 2	Model 3
Age at baseline	-0.078*** (0.019)	-0.053** (0.021)	-0.054*** (0.021)
Evaluation Time		0.006 (0.024)	0.025 (0.026)
MCI-worsened		-0.248*** (0.085)	-0.254** (0.113)
MCI-stable		-0.107 (0.058)	-0.044 (0.085)
Evaluation Time x MCI-worsened			0.025 (0.113)
Evaluation Time x MCI-stable			-0.071 (0.072)
Intercept	1.581*** (0.019)	1.610*** (0.032)	1.601*** (0.034)
Observations	624	624	624
Log Likelihood	-1,013.299	-1,007.706	-1,007.200
Akaike Inf. Crit.	2,032.597	2,027.413	2,030.400
Bayesian Inf. Crit.	2,045.570	2,053.359	2,064.994

Note: **p<0.05; ***p< 0.01

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es

Figure 1. Estimated marginal means and errors bars from Model 1 for PAL, PRM, DMS and SSP in the three groups across the three evaluation times.



Note: SE= Standard Error; BL = Baseline assessment; T1= Time 1 assessment; T2= Time 2 assessment.

Corresponding author: María Campos-Magdaleno, Department of Developmental Psychology, Rúa Xosé María Suárez Núñez, s/n. Campus Sur. 15782 Santiago de Compostela, (Galicia, Spain). Email: maria.campos@usc.es