CrossMark

RESEARCH ARTICLE

Learning to predict is spared in mild cognitive impairment due to Alzheimer's disease

Rosalind Baker[1] · Peter Bentham[2] · Zoe Kourtzi[3]

Abstract Learning the statistics of the environment is critical for predicting upcoming events. However, little is known about how we translate previous knowledge about scene regularities to sensory predictions. Here, we ask whether patients with mild cognitive impairment due to Alzheimer's disease (MCI–AD) that are known to have spared implicit but impaired explicit recognition memory are able to learn temporal regularities and predict upcoming events. We tested the ability of MCI–AD patients and age-matched controls to predict the orientation of a test stimulus following exposure to sequences of leftwards or rightwards oriented gratings. Our results demonstrate that exposure to temporal sequences without feedback facilitates the ability to predict an upcoming stimulus in both MCI–AD patients and controls. Further, we show that executive cognitive control may account for individual variability in predictive learning. That is, we observed significant positive correlations of performance in attentional and working memory tasks with post-training performance in the prediction task. Taken together, these results suggest a mediating role of circuits involved in cognitive control (i.e. frontal circuits) that may support the ability for predictive learning in MCI–AD.

Keywords Sequence learning · Sensory predictions · Attention · Memory

✉ Zoe Kourtzi
   zk240@cam.ac.uk

[1] School of Psychology, University of Birmingham, Birmingham B15 2TT, UK

[2] Birmingham and Solihull Mental Health Foundation Trust (BSMHFT), Edgbaston, Birmingham, UK

[3] Department of Psychology, University of Cambridge, Cambridge, UK

Introduction

Predicting upcoming events is critical for successful everyday interactions in complex environments from avoiding obstacles to forecasting the weather. It is thought that the brain achieves this challenge by taking into account information about the structure of the environment that is acquired through experience and training (Geisler 2008; Petrov et al. 2005). There is accumulating evidence that mere exposure to stimuli that co-occur in the environment facilitates our ability to extract spatial and temporal regularities (for reviews, see: Aslin and Newport 2012; Perruchet and Pacton 2006). In particular, observers report that structured combinations are more familiar than random contingencies after exposure to items (e.g. shapes, tones or syllables) that co-occur spatially or appear in a temporal sequence (Chun 2000; Fiser and Aslin 2002; Saffran et al. 1996, 1999; Turk-Browne et al. 2005). This statistical learning has been shown to facilitate object recognition (Brady and Chun 2007; Brady and Oliva 2008), language understanding (Misyak et al. 2010), social judgments (Kunda and Nisbett 1986) and inductive reasoning (Kemp and Tenenbaum 2009). This previous work suggests that observers acquire implicit knowledge of the regularities present in a scene, despite the fact that they may not be explicitly aware of its specific structure. However, little is known about how we translate this implicit knowledge of temporal structures to predictions of future events.

In our previous work (Baker et al. 2014), we have shown that exposure to temporal regularities in a scene facilitates observers to learn its global structure and use this implicitly acquired knowledge to predict upcoming sensory events. Neuroimaging studies have implicated the hippocampus and striatum in learning of probabilistic associations (Poldrack et al. 2001; Shohamy and Wagner 2008)

Springer

Fig. 1 Stimuli and design. Diagram illustrating the trial design: a first sequence of eight gratings was presented, followed by a second sequence that was interrupted by the presentation of a cue and test stimulus. The sequence continued after the participants indicated their response until all eight gratings were presented, indicating the end of the trial
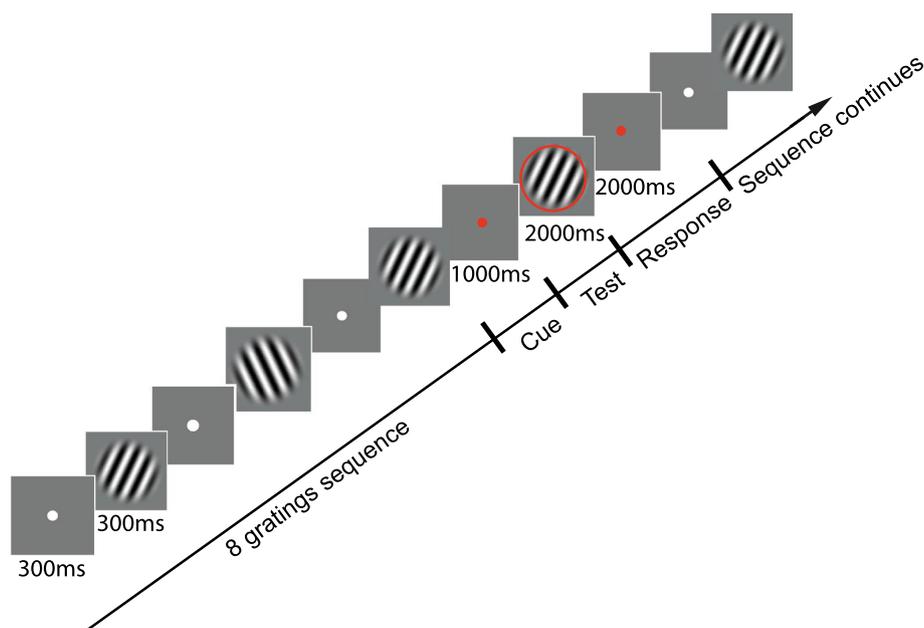
and temporal sequences (Gheysen et al. 2011; Hsieh et al. 2014; Rauch et al. 1997; Rose et al. 2011; Schapiro et al. 2012, 2014; Schendan et al. 2003).

Here, we ask whether the ability to acquire knowledge of predictive structures is maintained in patients with mild cognitive impairment due to Alzheimer's disease (MCI–AD). These patients are of particular interest, as they show explicit memory impairments (Hudon et al. 2006; Morris and Cummings 2005; Petersen et al. 1999) and hippocampal dysfunction (Bakker et al. 2012; Celone et al. 2006; Dickerson et al. 2004, 2005), but preserve their functional independence (Albert et al. 2011) and do not meet the clinical criteria for dementia. There is little evidence for learning of temporal structures in MCI–AD: while explicit temporal sequence learning is shown to require longer training in amnestic MCI–AD compared to age-matched controls, implicit temporal sequence learning is shown to be spared (Negash et al. 2007; Pirogovsky et al. 2013). Interestingly, previous work suggests that explicit learning is mediated by medial temporal lobe structures (e.g. hippocampus) that show dysfunction in MCI–AD, while implicit learning is mediated by striatal areas (Knowlton et al. 1996; Poldrack et al. 2001) that are spared in MCI–AD.

In the light of these previous findings, we test the hypothesis that preserved ability for implicit learning, despite hippocampal dysfunction, facilitates the ability of MCI–AD patients to predict upcoming sensory events after training on structured temporal sequences. We further test the hypothesis that preserved ability for learning of predictive structures may relate to cognitive capacity as indicated by attentional and working memory skills. To this end, we used a predictive learning task (Baker et al. 2014). In particular, we presented observers (MCI–AD

patients and age-matched controls) with a sequence of leftwards and rightwards oriented gratings that was interrupted by a test stimulus (Fig. 1). Observers had to maintain attention throughout the temporal sequence as the temporal position of the test stimulus was randomly chosen across trials and were asked to indicate whether the test stimulus matched their expectation or not. Participants were exposed to the sequences without feedback, facilitating implicit learning of the sequence structure, but they were asked to make an explicit judgment about the identity of the upcoming test stimulus. Thus, this task provides an explicit recognition measure of implicitly acquired knowledge, avoiding reaction time measurements that may be confounded by differences in speed of processing or response time between patients and controls. Our results demonstrate that the ability to predict the orientation of the test stimulus following exposure to structured sequences improved in both MCI–AD patients and controls. Further, we show that attentional and working memory skills may account for individual differences in task performance in both patients and controls. Taken together, these results suggest a mediating role of circuits involved in cognitive control (i.e. frontal circuits) in predicting sensory events based on previous knowledge about the environment's statistics.

Methods

Participants

Twenty-eight volunteers participated in this study (14 MCI–AD patients: 10 males, 4 females, mean age:

69.8 years; 14 age-matched controls: 10 males, 4 females, mean age: 67.7 years). The two groups did not differ significantly in their age ($t(26) = 0.133$, $p = 0.896$). All participants (patients and controls) were naïve to the aim of the study, had normal or corrected-to-normal vision (there were no differences in visual acuity between patients and controls) and gave written informed consent. This study was approved by the University of Birmingham Ethics Committee and the NHS National Research Ethics Committee, West Midlands. Patients, diagnosed with MCI–AD by their consultant psychiatrist, were recruited from the Birmingham and Solihull Memory Assessment and Advisory Service. Age-matched controls were recruited through advertising at the local community ($n = 11$) or were relatives of the MCI–AD patients who participated in the study ($n = 3$).

The diagnosis of MCI due to Alzheimer's disease was made by an experienced consultant psychiatrist (PB) using the National Institute on Ageing and Alzheimer's Association workgroup criteria (Albert et al. 2011) requiring: a deterioration in cognition reported by either the patient or a close informant; objective impairment in one or more cognitive domains (including memory, executive function, visuospatial skills, attention and language); preservation of independence in daily living activities; the absence of dementia; and an aetiology consistent with Alzheimer's disease pathophysiological process. Age-matched controls were screened using the Addenbrookes Cognitive Examination (ACE-III) (Hsieh et al. 2013). Scores for the controls (mean = 94.71; standard error = 0.86) were considered normal for the age of individual participants in comparison to the MCI–AD patients (mean = 86.5; standard error = 1.65), indicating lack of cognitive impairment for this group .

Experimental design

All participants (patients, controls) were tested in a set of cognitive tasks (working memory, selective and divided attention; see details for cognitive testing below) before starting training on the prediction task. Most participants ($n = 18$; patients = 11, controls = 7) completed 5 training sessions on the prediction task ($n = 8$; patients = 3, controls = 5 completed 4 sessions; $n = 2$; controls = 2 completed 3 sessions) depending on individuals' availability, with an average of 2.29 days between sessions (SD = 0.91).

Prediction task: stimuli

Stimuli comprised greyscale sinusoidal gratings that were presented at 10.8° visual angle, spatial frequency that ranged from 0.85° to 1 cycle per degree across trials, 100 % contrast and randomized phase. These gratings were rotated ±10° from vertical orientation (90°), resulting in gratings oriented at either 100° (left) or 80° (right). To avoid adaptation to the stimulus properties due to stimulus repetition, we randomized the phase and jittered the grating orientation within a range of −2° to 2° across trials. Stimuli were generated and presented using Psychtoolbox-3 (Brainard 1997; Pelli 1997). Stimuli were presented on a 21-inch CRT monitor (ViewSonic P225f 1280 × 1024 pixel, 85 Hz frame rate) at a distance of 45 cm.

We used these stimuli to generate two sequences, each comprising of eight gratings that were ordered, as shown below (1 refers to the leftwards oriented grating at −10°, and number 2 refers to the rightwards oriented grating at +10°):

Sequence A: 2 1 2 1 1 2 1 2
Sequence B: 1 1 2 1 2 2 1 2

Each grating orientation was presented four times in each sequence. Each sequence was repeated twice, resulting in 16 stimuli per trial. As all gratings were presented at the same rate, participants could not use stimulus duration to group elements together or segment the sequences. To ensure that the participants did not perform the task simply by memorizing the first or last stimuli in the sequence, the orientation of the first stimulus was randomized in each trial and the last three stimuli in each sequence were always the same. Finally, as the frequency of occurrence was matched for the two grating orientations in the sequence, participants were required to learn the order of the elements in the sequence (i.e. temporal order associations between pairs or triplets of oriented gratings).

Prediction task: design and training

For each trial, participants viewed 16 gratings (each sequence of eight gratings was repeated twice in a trial) presented sequentially on a grey background at the centre of the screen. Each grating was presented for 0.3 s followed by a fixation interval of 0.3 s. Participants were asked to respond to a test stimulus that appeared for up to 2000 ms surrounded by a red circle. The test stimulus was preceded by a cue (red dot presented for 1 s) and was followed by a white fixation dot (200 ms), which appeared as soon as the participant responded. Participants were instructed to respond (the maximum response time allowed was 2000 ms), indicating whether the test image had the same orientation (left vs. right) as the grating they expected to appear in that position in the sequence. The test stimulus appeared only in the second repeat of the sequence, and its position was randomized across trials. The test stimulus could appear in any position in the sequence except the

last three positions; stimuli in these positions were the same across trials. For each block, 50 % of the test stimuli were presented at the correct orientation for their position in the sequence. After the participant's response, the remaining gratings in the sequence were presented followed by a black cross (1 s), indicating the end of the sequence and the start of a new trial. There was no feedback across all sessions. In each training session, participants performed the prediction task for four blocks of 40 trials each (20 per sequence type) with a minimum 2-min break between blocks.

Prediction task: data analysis

We assessed behavioural performance by accuracy (per cent correct) across trials; that is, we computed whether the test grating was predicted correctly or not (i.e. the participants response matched the grating expected based on the presented sequence in each trial).

Cognitive testing

Ace-III

ACE-III is a clinical tool used to assess cognitive functioning, which takes approximately 25 min to administer. The Addenbrookes Cognitive Assessment (ACE) was developed as a brief test of cognitive function with the aim of early detection of dementia and differentiation into diagnostic subtypes (Mathuranath et al. 2000). The ACE was revised to improve its administration and sensitivity and has been shown to have very good psychometric properties (Mioshi et al. 2006). Due to copyright issues, some items required modification resulting in the ACE-III version which has been utilized in this study (Hsieh et al. 2013).

ACE-III has a maximum score of 100 and comprises of five subtests. *Attention*: orientation, three item registration, serial subtraction and recall questions are asked (18 points); *memory*: name and address learning, recall and recognition plus general knowledge questions are asked(26 points); *fluency*: the interviewee lists as many words in a category as they can within 1 min (14 points); *language:* images of objects and animals of varying familiarity are identified, words and phrases are repeated and two sentences are written about a recent event (26 points); and *visuospatial skills*: abstract letter perception, a clock drawn from memory and dot counting (16 points).

Memory: working memory task

The working memory task was designed following Luck and Vogel (1997). Coloured dots were displayed on a grey background for 500 ms, followed by a 1000 ms delay. After the delay, the dot display reappeared and one of the dots was highlighted by a white square. Participants reported whether the highlighted dot had remained the same colour on the second presentation. An initial display of two dots was used. By using a two-down one-up staircase and a step size of 1, we manipulated the number of dots in the display, resulting in 70.7 % performance. For example, each time the participant had two responses correct in a row, an additional dot would be added to the next trial's display, while for every incorrect answer, one dot was removed from the display for the next trial. Working memory thresholds (i.e. number of dots in the display) were calculated by averaging the last two-third reversals in each staircase. For each trial, each dot was randomly assigned a colour, and one dot was randomly chosen as the target. Each dot had a radius of 12 pixels, and dots were displayed in random locations within a 10 × 10 grid (jittered ± 10 pixels). Each block consisted of 10 staircase reversals, and participants completed three blocks, after which we computed the average threshold as their working memory score. In this task, a higher score (greater number of items in display) denotes better performance.

Attention: useful field of view

Useful Field of View (UFOV; Visual Awareness Inc.) is a task that assesses three attentional processes: processing speed, divided attention and selective attention (Edwards et al. 2005, 2006). The task has been validated by a test–retest reliability of 0.74. Each trial started with a fixation bounding box (1-s duration), followed by the test stimuli (variable duration between 16.7 and 500 ms; see below), a white noise visual mask to control for after images (1-s duration) and the response screen (displayed until a response was made). Participants responded using the mouse. The first test, 'processing speed', required participants to identify a centrally presented stimulus. This stimulus (a silhouette of 2 cm × 1.5 cm of a car or a truck) was presented on a black background inside a 3 cm × 3 cm white bounding box. Participants were asked to indicate whether the central stimulus comprised a car or truck by mouse click. The second task, 'divided attention', required participants to identify the central stimulus (car vs. truck) and also identify the location of a simultaneously presented peripheral stimulus (2 cm × 1.5 cm silhouette of a car). This peripheral stimulus was fixed at 11 cm from the central stimulus at one of the eight radial locations. The third task 'selective attention' followed the same procedure as 'divided attention', but the target stimuli were presented in the context of distractors (47 triangles of the same size and luminance as the targets). Participants were instructed to ignore the triangles and indicate whether the central stimulus comprised a car or a truck, as we all the location of the peripheral target. Using a double-staircase method, the duration of the display within each task varied between

16.7 and 500 ms. This allowed us to establish the minimal display duration at which the participant could correctly perform each of the three tests 75 % of the time. Thus, a lower score (shorter display duration) indicates better performance in these tasks.

Results

Performance on the prediction task improved for most participants (11/14 MCI–AD patients, 11/14 controls) as they gained more exposure to the temporal sequences. MCI–AD patients ($n = 3$) who did not improve during training showed mean performance of 40.4 % correct at the last training session. Healthy controls ($n = 3$) who did not improve during training showed mean performance of 48.3 % correct at the last training session. To quantify this learning effect, we compared the mean of the first two training blocks with the mean of the last two training blocks across participants (Fig. 2a), excluding participants who did not improve during training. To compare performance between MCI–AD patients and controls before and after training, we used a 2 (*session: pre- vs. post-test*) × 2 (*group: MCI–AD controls*) mixed design ANOVA. Our results showed that both MCI–AD patients and controls improved after training in the prediction task. That is, there was a main effect of session ($F(1,20) = 36.1$, $p < 0.001$) and significantly higher performance after than before training for both patients ($t(10) = -3.855$, $p = 0.003$) and controls ($t(10) = -4.846$, $p = 0.001$). Further, controls showed overall higher performance than patients as indicated by a significant main effect of group ($F(1,20) = 6.32$, $p = 0.021$). In particular, higher performance for controls was significant before ($t(20) = -2.479$, $p = 0.022$) but only marginally after ($t(20) = -1.94$, $p = 0.067$) training. Interestingly, both patients and controls improved similarly after training as indicated by a non-significant Session × Group interaction ($F(1,20) = 0.1$, $p = 0.921$). This was confirmed by an additional analysis (Fig. 2b) that showed no significant difference ($t(20) = 0.101$, $p = 0.921$) in behavioural improvement between groups, as calculated by subtracting performance in the last two training blocks from performance in the first two training blocks (Fig. 2b).

To control for possible differences in learning-dependent improvements due to differences in the numbers of training sessions across participants, we conducted two additional analyses. First, we considered all participants who had at least four training sessions (11 patients and 9 controls), excluding participants (3 patients and 2 controls) who did not improve during training. This analysis showed the same pattern of results as in Fig. 2; that is, we observed a significant main effect of session ($F(1,18) = 25.0$, $p < 0.001$), but no significant interaction between session and group ($F(1,18) = 0.752$,
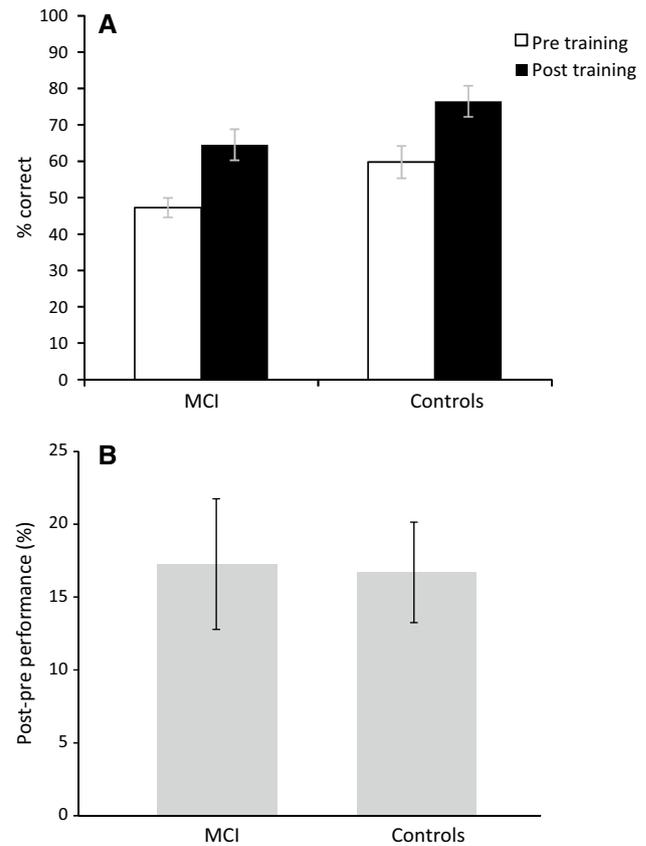

Fig. 2 Training on prediction task. **a** Mean (across participants) per cent correct performance is shown for the start (mean of the first two training blocks of the first session) and the end (mean of the last two training blocks of the last session) of training for patients and controls. **b** Mean improvement for MCI–AD patients and controls calculated by subtracting performance at the end from performance at the start of training. *Error bars* indicate standard error of the mean

$p = 0.397$), suggesting similar improvement for patients and controls. Second, we conducted the same analyses using the data from the third session as the final session for all participants (11 patients and 11 controls). This analysis showed the same pattern of results as in Fig. 2; that is, we observed a significant main effect of session ($F(1,20) = 18.18$, $p < 0.001$), but no significant interaction between session and group ($F(1,20) = 1.046$, $p = 0.319$). Further comparisons showed a significant difference in performance between the first and the third sessions ($t(25) = -4.35$, $p < 0.001$), but not between the third and the last sessions ($t(25) = -1.35$, $p = 0.188$), suggesting that the first three training sessions provided adequate information to capture learning improvement in both young and older participants.

Relating cognitive abilities to learning performance

We then asked whether cognitive control abilities (i.e. attention, working memory) relate to learning improvement

in the prediction task. Although MCI–AD patients and controls did not differ significantly in processing speed ($t(20) = 1.306$, $p = 0.206$), MCI–AD patients showed lower performance in selective attention ($t(17) = 2.952$, $p = 0.009$), divided attention ($t(20) = 3.474$, $p = 0.002$) and working memory ($t(20) = -2.377$, $p = 0.028$) tasks. This is consistent with the assessment of the patients based on ACE-III that tests similar cognitive abilities and showed lower scores for MCI–AD patients compared to controls ($t(20) = -3.78$, $p < 0.001$). Interestingly, we observed individual variability in cognitive testing and predictive learning scores in both patients and controls. That is, performance in the prediction task after training ranged from 45 to 91.25 % correct for MCI–AD patients and from 53.75 to 93.75 % correct for controls. Performance in selective attention ranged from 156.7 to 500 ms for MCI–AD patients and from 30.1 to 183.3 ms for controls; performance in divided attention ranged from 16.7 to 296.7 ms for MCI–AD patients and from 16.7 to 133.4 ms for controls; performance in working memory ranged from 1.17 to 5.16 number of dots for MCI–AD patients and from 2.05 to 5.16 number of dots for controls.

To further investigate whether individual variability in cognitive abilities relates to performance in predictive learning, we correlated scores in these tasks collected before training with performance in the prediction task after training (Fig. 3). In particular, we observed a significant correlation between working memory scores ($r = 0.665$, $N = 28$, $p < 0.001$) and divided attention scores ($r = -0.463$, $p = 0.009$; $N = 26$: missing data for two participants due to technical problems) with post-training performance in the prediction task across all participants (i.e. patients and controls). Further, scores in the divided attention task showed significant correlations with behavioural improvement in the prediction task (that is difference in performance between sessions) for both patients ($R = -0.468$, $p = 0.05$) and controls ($R = -0.468$, $p = 0.05$). We also observed a marginally significant correlation between selective attention scores and post-training performance in the prediction task ($r = -0.341$, $p = 0.056$; $N = 23$: missing data for five participants due to technical problems). The negative correlations for the divided and selective attention are due to faster display times (i.e. lower values), indicating better attentional performance.

Further, we conducted the same correlations with the data from participants who completed at least four sessions. This analysis showed a significant correlation between performance in the fourth session and working memory ($R = 0.377$, $p = 0.05$), divided attention ($R = -0.425$, $p = 0.031$) and selective attention ($R = -0.408$, $p = 0.05$). In addition, the same correlational analyses using the data from the third session as the final session for all participants showed significant correlations with working memory

($R = 0.447$, $p = 0.017$), selective attention ($R = -0.405$, $p = 0.055$) and marginally divided attention ($R = -0.367$, $p = 0.065$). Finally, performance in the fourth ($R = 0.473$, $p = 0.018$) or third ($R = 0.850$, $p < 0.001$) session was significantly correlated with performance in the final session, suggesting that learning-dependent improvement was not confounded by differences in the numbers of training sessions across participants.

Separate correlations for MCI–AD patients and controls showed similar trends with correlations across all participants (Fig. 3); however, these correlations need to be interpreted cautiously due to the smaller number of participants per group. In particular, for controls we observed significant correlations for all tasks (WM: $R = 0.603$, $p = 0.011$; DA: $R = -0.824$, $p < 0.001$, SA: $R = -0.661$, $p = 0.013$). For patients, we observed similar correlations that reached significance for working memory ($R = 0.504$, $p = 0.033$) but not for the attentional tasks (DA: $R = -0.066$, $p = 0.41$, SA: $R = -0.109$, $p = 0.35$). Fisher's z test did not show any significant differences in the correlations between groups for working memory ($z = 0.34$, $p = 0.734$), or selective attention ($z = -1.61$, $p = 0.107$), consistent with similar trends between patients and controls. Taken together, these analyses suggest similar patterns of correlations for patients and controls, although correlations for patients were weaker possibly due to higher variability between participants in this group.

These results were confirmed by a multiple regression analysis, showing that attention and working memory explain significantly ($F(3,18) = 4.19$, $p = 0.024$) 45.6 % of the variance in performance in the prediction task ($R = 0.347$). Similar analysis conducted separately for patients and controls showed significant results for controls ($F(3.9) = 5.200$, $p = 0.02$) and a marginal effect for patients ($F(3.11) = 2.366$, $p = 0.07$). Taken together, these results suggest that participants (MCI–AD patients and controls) with better attentional and working memory skills are more likely to improve in predicting future events following training on temporal sequences.

Discussion

Our results demonstrate that exposure to temporal sequences without feedback facilitates the ability of MCI–AD patients to predict upcoming events. Both patients and controls showed similar improvement after training in the prediction task that correlated with performance in working memory and attention tasks, suggesting a role of cognitive control abilities in learning to predict sensory events based on previous knowledge of temporal regularities.

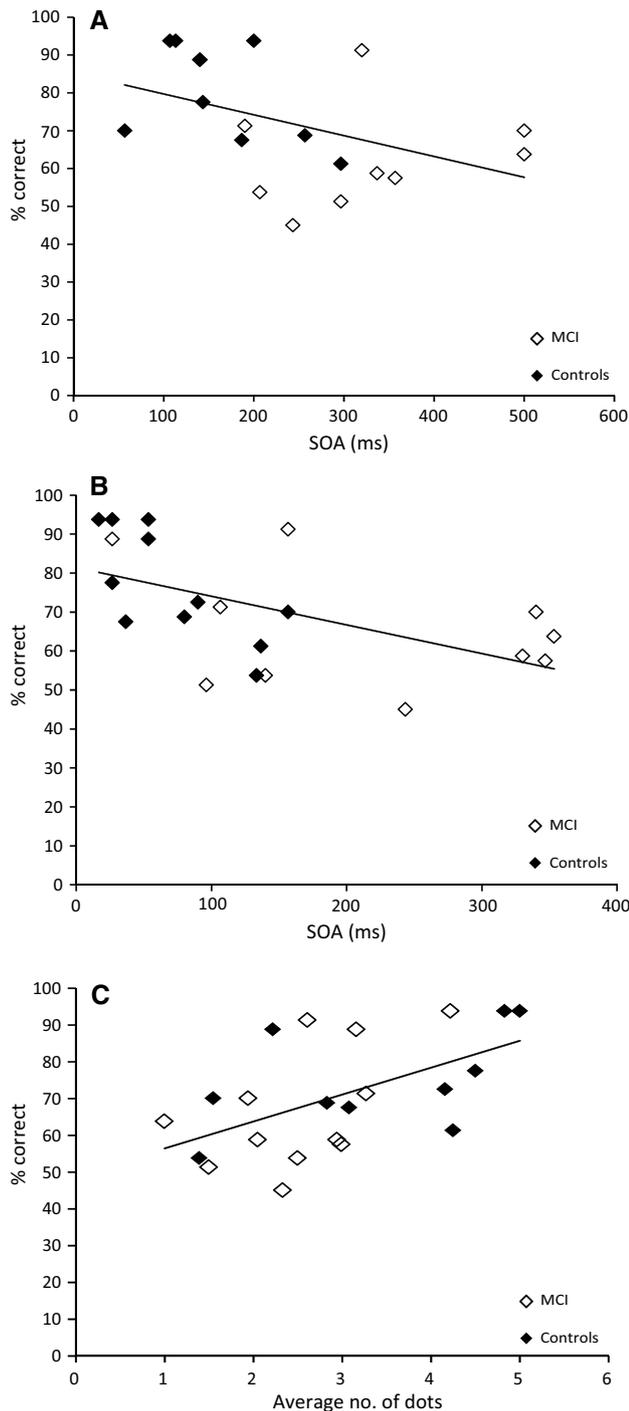Consistent with our previous behavioural work (Baker et al. 2014), we demonstrate that MCI–AD patients

tasks: visuomotor sequence learning (Nissen and Bullemer 1987), artificial grammar learning (Reber 1967), probabilistic category learning (Knowlton et al. 1994) and contextual cue learning (Chun and Jiang 1998). In our study, participants were exposed to the sequences without feedback, but were asked to make an explicit judgment about the identity of the upcoming test stimulus (leftwards vs. rightwards oriented grating), making them aware of the dependencies between the stimuli presented in the sequence. However, our experimental design makes it unlikely that the participants memorized specific item positions or the full sequences. Further, debriefing the participants showed that it was unlikely that the participants explicitly memorized the sequences. In particular, participants could not freely recall the sequences after training or correctly indicate the number of trained sequences.

Our study tests the role of sequence learning on explicit predictive judgments related to visual recognition in MCI–AD patients. Previous work on learning temporal sequences has focused on implicit measures of sequence learning, such as familiarity judgments or reaction times. For example, the Serial Reaction Time Task (Nissen and Bullemer 1987); for review, see (Schwarb and Schumacher 2012) involves participants learning visuomotor associations between spatial locations on a computer screen and response keys; locations on the screen are activated following a predetermined sequence, and participants are asked to press the corresponding keys. Training results in faster reaction times for trained than random sequences. However, using reaction times as a measure of anticipation of upcoming events may be problematic with patients and older adults that show generally reduced speed of processing and longer response times (Curran 1997; Simon et al. 2012). In contrast, using an explicit prediction test, we demonstrate that predictions related to identification of objects are facilitated by implicit knowledge of temporal context.

Our findings are consistent with previous work, suggesting that MCI patients are not impaired in implicit learning

accumulate information about temporal regularities through repeated exposure to an environment and predict future events. Although we used deterministic sequences, we ensured that observers learned the global sequence structure (i.e. temporal order statistics across items rather than temporal item positions in the sequence) by matching the frequency of occurrence of each item (i.e. grating orientation) in the sequence. Previous studies have suggested that learning of regularities may occur implicitly in a range of

tasks (Negash et al. 2007), while explicit learning may require longer training periods for MCI patients (Pirogovsky et al. 2013). This is in accordance with studies proposing that explicit learning is mediated by medial temporal lobe structures (e.g. hippocampus) that show dysfunction in MCI, while implicit learning is mediated by a fronto-striatal network that is spared in MCI. Interestingly, our study shows that patients and controls with better attentional and working memory skills show better performance after training on the prediction task. These skills are thought to implicate frontal circuits (Corbetta and Shulman 2002) that when damaged are shown to impair performance in tasks that involve updating strategies for future predictions (Danckert et al. 2012). Taken together, these findings suggest a mediating role of frontal circuits in MCI–AD that may facilitate learning and compensate against hippocampal dysfunction.

Finally, predicting conversion rate from MCI to Alzheimer's disease is a key question in clinical neuroscience. 14–18 % of those aged over 70 years meet the criteria for MCI, and patients are likely to develop dementia, in the order of 10–15 % per annum (Petersen et al. 2009). Future work including larger numbers of patients and follow-ups would allow us to test whether this prediction task could serve as a diagnostic tool and/or form the basis of a rehabilitative training programme. Although in this study we did not test long-term effects of training, our previous studies (Baker et al. 2014) have shown that improvement in the prediction task lasted for a prolonged period (up to 3 months), suggesting that training resulted in consolidated knowledge of the sequence. Future work is needed to investigate whether longer-term training may result in stronger and longer-lasting improvement following training on the prediction task. Further brain imaging work testing for compensatory involvement of frontal circuits consistent with improved cognitive control skills after training on the prediction task will advance our understanding about the neural mechanisms that may support training of cognitive abilities in MCI against the progression of cognitive decline.

References

Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, Phelps CH (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7(3):270–279. doi:10.1016/j.jalz.2011.03.008

Aslin RN, Newport EL (2012) Statistical learning: from acquiring specific items to forming general rules. Curr Dir Psychol Sci 21(3):170–176. doi:10.1177/0963721412436806

Baker R, Dexter M, Hardwicke TE, Goldstone A, Kourtzi Z (2014) Learning to predict: exposure to temporal sequences facilitates prediction of future events. Vision Res 99:124–133. doi:10.1016/j.visres.2013.10.017

Bakker A, Krauss GL, Albert MS, Speck CL, Jones LR, Stark CE, Gallagher M (2012) Reduction of hippocampal hyperactivity improves cognition in amnestic mild cognitive impairment [randomized controlled trial research support, N.I.H., extramural]. Neuron 74(3):467–474. doi:10.1016/j.neuron.2012.03.023

Brady TF, Chun MM (2007) Spatial constraints on learning in visual search: modeling contextual cuing. J Exp Psychol Hum Percept Perform 33(4):798–815. doi:10.1037/0096-1523.33.4.798

Brady TF, Oliva A (2008) Statistical learning using real-world scenes—extracting categorical regularities without conscious intent. Psychol Sci 19(7):678–685. doi:10.1111/j.1467-9280.2008.02142.x

Brainard DH (1997) The psychophysics toolbox. Spat Vis 10(4):433–436

Celone KA, Calhoun VD, Dickerson BC, Atri A, Chua EF, Miller SL, Sperling RA (2006) Alterations in memory networks in mild cognitive impairment and Alzheimer's disease: an independent component analysis. J Neurosci 26(40):10222–10231. doi:10.1523/jneurosci.2250-06.2006

Chun MM (2000) Contextual cueing of visual attention. Trend Cogn Sci 4(5):170–178. doi:10.1016/s1364-6613(00)01476-5

Chun MM, Jiang YH (1998) Contextual cueing: implicit learning and memory of visual context guides spatial attention. Cogn Psychol 36(1):28–71. doi:10.1006/cogp.1998.0681

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3(3):201–215. doi:10.1038/nrn755

Curran T (1997) Effects of aging on implicit sequence learning: accounting for sequence structure and explicit knowledge. Psychol Res 60:24–41

Danckert J, Stoettinger E, Quehl N, Anderson B (2012) Right hemisphere brain damage impairs strategy updating. Cereb Cortex 22(12):2745–2760. doi:10.1093/cercor/bhr351

Dickerson BC, Salat DH, Bates JF, Atiya M, Killiany RJ, Greve DN, Sperling RA (2004) Medial temporal lobe function and structure in mild cognitive impairment [research support, Non-U.S. Gov't research support, U.S. Gov't, P.H.S.]. Ann Neurol 56(1):27–35. doi:10.1002/ana.20163

Dickerson BC, Salat DH, Greve DN, Chua EF, Rand-Giovannetti E, Rentz DM, Sperling RA (2005) Increased hippocampal activation in mild cognitive impairment compared to normal aging and AD. Neurology 65(3):404–411. doi:10.1212/01.wnl.0000171450.97464.49

Edwards JD, Vance DE, Wadley VG, Cissell GM, Roenker D, Ball KK (2005) Reliability and validity of useful field of view test scores as administered by personal computer. J Clin Exp Neuropsychol 27(5):529–543. doi:10.1080/13803390490515432

Edwards JD, Ross LA, Wadley VG, Clay OJ, Crowe M, Roenker DL, Ball KK (2006) The useful field of view test: normative data for older adults. Arch Clin Neuropsychol 21(4):275–286. doi:10.1016/j.acn.2006.03.001

Fiser J, Aslin RN (2002) Statistical learning of higher-order temporal structure from visual shape sequences. J Exp Psychol Learn Mem Cogn 28(3):458–467. doi:10.1037//0278-7393.28.3.458

Geisler WS (2008) Visual perception and the statistical properties of natural scenes. Ann Rev Psychol 59:167–192

Gheysen F, Van Opstal F, Roggeman C, Van Waelvelde H, Fias W (2011) The neural basis of implicit perceptual sequence learning. Front Hum Neurosci 5:137. doi:10.3389/fnhum.2011.00137

Hsieh S, Schubert S, Hoon C, Mioshi E, Hodges JR (2013) Validation of the Addenbrooke's cognitive examination III in frontotemporal dementia and Alzheimer's disease. Dement Geriatr Cogn Disord 36(3–4):242–250. doi:10.1159/000351671

Hsieh LT, Gruber MJ, Jenkins LJ, Ranganath C (2014) Hippocampal activity patterns carry information about objects in temporal context. Neuron 81(5):1165–1178. doi:10.1016/j.neuron.2014.01.015

Hudon C, Belleville S, Souchay C, Gely-Nargeot M-C, Chertkow H, Gauthier S (2006) Memory for gist and detail information in Alzheimer's disease and mild cognitive impairment. Neuropsychology 20(5):566–577. doi:10.1037/0894-4105.20.5.566

Kemp C, Tenenbaum JB (2009) Structured statistical models of inductive reasoning (vol 116, pg 20, 2009). Psychol Rev 116(2):461. doi:10.1037/a0015514

Knowlton BJ, Squire LR, Gluck MA (1994) Probabilistic classification learning in amnesia. Learn Mem 1(2):106–120

Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. Science 273(5280):1399–1402. doi:10.1126/science.273.5280.1399

Kunda Z, Nisbett RE (1986) The psychometrics of everyday life. Cogn Psychol 18(2):195–224. doi:10.1016/0010-0285(86)90012-5

Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. Nature 390:279–281

Mathuranath PS, Nestor PJ, Berrios GE, Rakowicz W, Hodges JR (2000) A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. Neurology 55(11):1613–1620

Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR (2006) The Addenbrooke's cognitive examination revised (ACE-R): a brief cognitive test battery for dementia screening. Int J Geriatr Psychiatry 21(11):1078–1085. doi:10.1002/gps.1610

Misyak JB, Christiansen MH, Tomblin JB (2010) Sequential expectations: the role of prediction-based learning in language. Top Cogn Sci 2(1):138–153. doi:10.1111/j.1756-8765.2009.01072.x

Morris JC, Cummings J (2005) Mild cognitive impairment (MCI) represents early-stage Alzheimer's disease. J Alzheimers Dis 7(3):235–239

Negash S, Petersen LE, Geda YE, Knopman DS, Boeve BF, Smith GE, Petersen RC (2007) Effects of ApoE genotype and mild cognitive impairment on implicit learning. [comparative study research support, N.I.H., extramural research support, Non-U.S. Gov't]. Neurobiol Aging 28(6):885–893. doi:10.1016/j.neurobiolaging.2006.04.004

Nissen MJ, Bullemer P (1987) Attentional requirements of learning—evidence from performance measures. Cogn Psychol 19(1):1–32. doi:10.1016/0010-0285(87)90002-8

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10(4):437–442

Perruchet P, Pacton S (2006) Implicit learning and statistical learning: one phenomenon, two approaches. Trends Cogn Sci 10(5):233–238. doi:10.1016/j.tics.2006.03.006

Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E (1999) Mild cognitive impairment—clinical characterization and outcome. Arch Neurol 56(3):303–308. doi:10.1001/archneur.56.3.303

Petersen RC, Roberts RO, Knopman DS, Boeve BF, Geda YE, Ivnik RJ, Jack CR Jr (2009) Mild cognitive impairment: ten years later. Neurol Rev 66(12):1447

Petrov AA, Dosher BA, Lu ZL (2005) The dynamics of perceptual learning: an incremental reweighting model. Psychol Rev 112(4):715–743. doi:10.1037/0033-295x.112.4.715

Pirogovsky E, Holden HM, Jenkins C, Peavy GM, Salmon DP, Galasko DR, Gilbert PE (2013) Temporal sequence learning in healthy aging and amnestic mild cognitive impairment. Exp Aging Res 39(4):371–381. doi:10.1080/0361073x.2013.808122

Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Creso Moyano J, Myers C, Gluck MA (2001) Interactive memory systems in the human brain [research support, Non-U.S. Gov't]. Nature 414(6863):546–550. doi:10.1038/35107080

Rauch SL, Whalen PJ, Savage CR, Curran T, Kendrick A, Brown HD, Rosen BR (1997) Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. Hum Brain Mapp 5(2):124–132

Reber AS (1967) Implicit learning of artificial grammars. J Verbal Learn Verbal Behav 6(6):855. doi:10.1016/s0022-5371(67)80149-x

Rose M, Haider H, Salari N, Buchel C (2011) Functional dissociation of hippocampal mechanism during implicit learning based on the domain of associations. J Neurosci 31(39):13739–13745. doi:10.1523/JNEUROSCI.3020-11.2011

Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. Science 274(5294):1926–1928

Saffran JR, Johnson EK, Aslin RN, Newport EL (1999) Statistical learning of tone sequences by human infants and adults. Cognition 70(1):27–52. doi:10.1016/s0010-0277(98)00075-4

Schapiro AC, Kustner LV, Turk-Browne NB (2012) Shaping of object representations in the human medial temporal lobe based on temporal regularities. Curr Biol 22(17):1622–1627. doi:10.1016/j.cub.2012.06.056

Schapiro AC, Gregory E, Landau B, McCloskey M, Turk-Browne NB (2014) The necessity of the medial temporal lobe for statistical learning. J Cogn Neurosci. doi:10.1162/jocn_a_00578

Schendan HE, Searl MM, Melrose RJ, Stern CE (2003) An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron 37(6):1013–1025

Schwarb H, Schumacher EH (2012) Generalized lessons about sequence learning from the study of the serial reaction time task. Adv Cogn Psychol Univ Financ Manag Warsaw 8(2):165–178. doi:10.2478/v10053-008-0113-1

Shohamy D, Wagner AD (2008) Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events [research support, N.I.H., extramural research support, Non-U.S. Gov't]. Neuron 60(2):378–389. doi:10.1016/j.neuron.2008.09.023

Simon SS, Yokomizo JE, Bottino CM (2012) Cognitive intervention in amnestic Mild Cognitive Impairment: a systematic review. [Review]. Neurosci Biobehav Rev 36(4):1163–1178. doi:10.1016/j.neubiorev.2012.01.007

Turk-Browne NB, Junge JA, Scholl BJ (2005) The automaticity of visual statistical learning. J Exp Psychol General 134(4):552–564. doi:10.1037/0096-3445.134.4.552