Highlights

- Learning temporal structures facilitates prediction of sensory events in MCI-AD
- Dissociable brain circuits for predictive learning in MCI-AD vs. healthy controls
- A cortico-striatal-cerebellar network mediates predictive learning in MCI-AD

Learning temporal statistics for sensory predictions in mild cognitive impairment

Caroline Di Bernardi Luft[1], Rosalind Baker[2], Peter Bentham[3], Zoe Kourtzi[4]

[1]Department of Psychology, Goldsmiths, University of London, London UK
[2]School of Psychology, University of Birmingham, Birmingham, B15 2TT, UK
[3] Birmingham and Solihull Mental Health Foundation Trust (BSMHFT), Edgbaston, Birmingham, UK
[4]Department of Psychology, University of Cambridge, Cambridge UK

Correspondence:
Zoe Kourtzi
Department of Psychology
University of Cambridge
Cambridge, UK
Email: zk240@cam.ac.uk

Keywords: sequence learning, sensory predictions, fMRI

Running title: predictive learning in MCI-AD

Acknowledgements: We would like to thank Matthew Dexter for help with software development and the Birmingham and Solihull Mental Health Foundation Trust Memory Assessment Service for assistance with patient recruitment and screening.

Funding: This work was supported by grants to PB from Birmingham and Solihull Mental Health Foundation Trust Research and Development, and to ZK from the Leverhulme Trust [RF-2011-378] and the [European Community's] Seventh Framework Programme [FP7/2007-2013] under agreement PITN-GA-2011-290011.

Abstract

Training is known to improve performance in a variety of perceptual and cognitive skills. However, there is accumulating evidence that mere exposure (i.e. without supervised training) to regularities (i.e. patterns that co-occur in the environment) facilitates our ability to learn contingencies that allow us to interpret the current scene and make predictions about future events. Recent neuroimaging studies have implicated fronto-striatal and medial temporal lobe brain regions in the learning of spatial and temporal statistics. Here, we ask whether patients with mild cognitive impairment due to Alzheimer's disease (MCI-AD) that are characterized by hippocampal dysfunction 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 orientated 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. However, our fMRI results demonstrate that MCI-AD patients recruit an alternate circuit to hippocampus to succeed in learning of predictive structures. In particular, we observed stronger learning-dependent activations for structured sequences in frontal, subcortical and cerebellar regions for patients compared to age-matched controls. Thus, our findings suggest a cortico-striatal-cerebellar network that may mediate the ability for predictive learning despite hippocampal dysfunction in MCI-AD.

Introduction

Learning through supervised and extensive training is known to shape perceptual and cognitive skills in the adult human brain (Draganski et al., 2004; Goldstone, 1998; Kourtzi, 2010). However, there is accumulating evidence that mere exposure (i.e. without feedback) to stimuli that co-occur in the environment facilitates our ability to learn contingencies and extract spatial and temporal regularities (for reviews see: (Aslin & Newport, 2012; Perruchet & 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 & Aslin, 2002a; Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999; Turk-Browne, Junge, & Scholl, 2005). This statistical learning has been shown to facilitate object recognition (Brady & Chun, 2007; Brady & Oliva, 2008), language understanding (Misyak, Christiansen, & Tomblin, 2010), social judgments (Kunda & Nisbett, 1986) and inductive reasoning (Kemp & 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.

In our previous work (Baker, Dexter, Hardwicke, Goldstone, & Kourtzi, 2014), we have shown that exposure to temporal regularities in a scene facilitates observers to learn its global structure and use this knowledge to predict upcoming sensory events. Recent neuroimaging studies have implicated fronto-striatal and medial temporal lobe regions in the learning of temporal statistics. In particular, the striatum and the hippocampus have been implicated in learning of probabilistic associations (Poldrack et al., 2001; Shohamy & Wagner, 2008) and temporal sequences (Gheysen, Van Opstal, Roggeman, Van Waelvelde, & Fias, 2011; Hsieh, Gruber, Jenkins, & Ranganath, 2014; Rauch et al., 1997; Rose, Haider, Salari, & Buchel,

2011; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014; Schapiro, Kustner, & Turk-Browne, 2012; Schendan, Searl, Melrose, & Stern, 2003a).

Here, we ask whether patients with mild cognitive impairment due to Alzheimer's disease (MCI-AD) are able to exploit temporal regularities to predict upcoming events. MCI-AD patients are of particular interest as they show memory impairments (especially in episodic memory tasks) (Hudon et al., 2006; Morris & Cummings, 2005; Petersen et al., 1999) and hippocampal dysfunction (Bakker et al., 2012; Celone et al., 2006; Dickerson et al., 2004; Dickerson, Salat, Greve, Chua, et al., 2005), but preserve their functional independence (Albert et al., 2011) and do not to meet the clinical criteria for dementia. Previous work suggests that MCI patients are not impaired in implicit temporal sequence learning, while explicit temporal sequence learning is shown to require longer training in amnestic MCI-AD compared to age-matched controls (Pirogovsky et al., 2013). However, the brain circuits that may support implicit learning of temporal structures in MCI-AD remain largely unknown. Here, using fMRI (functional magnetic resonance imaging) we test for alternate brain circuits that may support implicit learning of temporal regularities despite hippocampal dysfunction in MCI-AD. To this end, we use a predictive learning task (Baker et al., 2014), that allows us to test for brain circuits that support explicit predictions based on implicitly acquired knowledge.

In particular, we presented MCI-AD patients and age-matched controls with a sequence of leftwards and rightwards oriented gratings that was interrupted by a test stimulus (Figure 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 orientation of the test stimulus matched the expected stimulus or not. 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 behavioral results show 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, our fMRI results provide evidence for a cortico-striatal-cerebellar network that may facilitate learning of predictive structures despite hippocampal dysfunction in MCI-AD.

FIGURE 1 ABOUT HERE

Materials and Methods

Participants

Twenty-one volunteers (11 MCI patients, 10 age matched controls) participated in this study. The data from two patients and one control were excluded from further analysis due to excessive head movement; therefore data from nine patients (7 male and 2 female; mean age 69.8 years; range 53 – 86 years) and nine controls (6 male and 3 female; mean age 65.1 years; range 56 – 83 years) were considered for further analysis. There was no significant age difference between groups ($t(16) = 1.02$, $p = .321$). All participants were naïve to the aim of the study, had normal or corrected-to-normal vision 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 or were relatives of the MCI patients that 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 Aging 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; absence of dementia and an aetiology consistent with Alzheimer's disease pathophysiological process. No patients with vascular-related disease were included in the study. Age-matched controls were screened using the Addenbrookes Cognitive Examination (ACE-III) (Dickerson et al., 2004). Scores for controls (mean = 96.0; standard error =0.7) compared to MCI-AD patients (mean = 87.6; standard error = 1.38) were considered normal for the age of individual participants, indicating lack of cognitive impairment for this group.

Stimuli

Stimuli comprised grayscale sinusoidal gratings that were presented at 10.8^{o} visual angle, spatial frequency that ranged from 0.85 to 1^{o} across trials, 100% contrast and randomized phase. These gratings were rotated +/- 10^{o} from vertical orientation (90^{o}), resulting in gratings oriented at either 100^{o} (left) or 80^{o} (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^{o} to 2^{o} across trials.

We used these stimuli to generate two sequences, each comprising of 8 gratings that were ordered, as shown below (1 refers to the leftwards oriented grating at -10 degrees and number 2 refers to the rightwards oriented grating at +10 degrees):

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. Further, to ensure that participants did not learn the task simply by memorizing the last orientations in the sequence, the last three stimuli were the same across all sequences. These manipulations preserved equal frequency of appearance for the two orientations across trials. Finally, as the frequency of occurrence was matched for the two grating orientations in the sequence and the participants did not know how many items each sequence contained, to perform the task participants were required to learn the order of the elements in the sequence (e.g. temporal order associations among pairs or triplets of oriented gratings). In addition to these structured sequences (A and B), random sequences (comprising 16 gratings presented in random order) were generated for the scanning sessions. A different random sequence was generated for each trial, so there was no repetition of the random sequences.

Stimuli were generated and presented using Psychtoolbox-3 (Brainard, 1997; Pelli, 1997). For the behavioural training sessions, stimuli were presented on a 21-inch CRT monitor (ViewSonic P225f 1280 x1024 pixel, 85 Hz frame rate) at a distance of 45 cm. For the pre and post-test fMRI scans, stimuli were presented using a projector and a mirror set-up (1280 x 1024 pixel, 60 Hz frame rate) at viewing distance of 67.5 cm. In order to keep the same visual angle for both training and scanning sessions (10.8), the stimulus size was adjusted according to the viewing distance.

Experimental Design

Participants took part in two fMRI scans (pre- and post-training) before and after behavioral training in the lab. Participants completed 3-5 training sessions depending on individuals' availability, with an average of 2.95 days between training sessions for the MCI group (standard error = 0.41) and 3.03 days for the control group (standard error = 0.71). The post-training scan took place the day after the last behavioral training session. Most participants (n =14: MCI= 8; controls=6) completed 5 training sessions. Three participants (2 controls, 1 patient) completed 4 sessions, and one control participant completed 3 sessions.

Behavioural Training

For each trial, participants viewed 16 gratings (each sequence of 8 gratings was repeated twice in a trial) presented sequentially on a grey background at the centre of the screen. All stimuli were presented at the same rate; that is, 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 0.3 ms surrounded by a red circle (0.3 s). The test stimulus was preceded by a cue (red dot presented for 1 s) and was followed by a white fixation (1700 ms). 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 pseudo randomized across trials. The test stimulus could appear in any of the following positions: 9,10,11,12,13 but not the last three positions; stimuli in these positions were the same across trials. For each run, 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 4 runs of 30 trials each (15 per sequence type) with a minimum two-minute break between runs.

fMRI design

Participants completed 5 to 8 runs of the prediction task without feedback in each scanning session (Fig. 1). Each run comprised 5 blocks of structured and 5 blocks of random sequences (3 trials per block; 1 sequence per trial) presented in a random counterbalanced order. The same number of A and B sequences were presented in each run and the sequence order was

randomized across the run. Each trial lasted for 10 seconds, resulting in 30 s long blocks. After each stimulus block, a white fixation was presented for 10 s. Each run started and ended with a fixation block. Similar to the task used for behavioral training, in each trial a sequence (structured) of 8 gratings was presented twice. For random sequence trials, 16 left or right oriented gratings were presented in a random order. Each grating was presented for 0.25 s followed by fixation for 0.2 s. In each trial, participants were presented with a test stimulus that appeared for 0.25 ms surrounded by a red circle (0.25 s) and preceded by a cue (red dot, 0.75 s). The test grating appeared only in the second repeat of the sequence and its position was pseudo randomized across trials. The test stimulus could appear in any of the following positions: 9,10,11,12,13 but not the last three positions; stimuli in these positions were the same across trials. After the test stimulus, a fixation dot was presented for 2 s (0.2 as red, remaining 1.8 s as white), instructing the participants to respond whether the test matched (button 1) the predicted stimulus or not (button 2). In half of the trials, the test image was consistent with the sequence ('correct' response required). After the response, the remaining gratings were presented till the end of the sequence, followed by an 'X' cue (0.65 s) marking the end of the trial.

fMRI Data Acquisition

fMRI data were acquired in a 3T Achieva Philips scanner at the Birmingham University Imaging Centre using an thirty two-channel head coil. Anatomical images were obtained using a sagittal three-dimensional T1-weighted sequence (voxel size=1 × 1 × 1 mm, slices=175) for localization and visualization of functional data. Functional data were acquired with a T2*weighted EPI sequence with 32 slices (whole-brain coverage; TR 2 s; TE 35 ms; flip angle 73; resolution 2.5*2.5*4 mm).

Eye movement recordings

We recorded eye-movements using the ASL 6000 Eye-tracker (Applied Science Laboratories, Bedford, MA, sampling rate: 60Hz) in the scanner. Eye tracking data were pre-processed using the EyeNal Data Analysis software (Applied Science Laboratories, Bedford, MA) and analyzed using custom toolbox based on Matlab (Mathworks, MA) software. Eye-tracking with this infra-red system could not be completed successfully for participants that wore glasses due to disruptive reflections. We report eye movement data from participants (n=6, 3 patients, 3 controls) for whom the eye-tracker could be calibrated successfully within the time limits of the scanning session. We computed (A) horizontal eye position, (B) vertical eye position, (C) proportion of saccades for each condition at different saccade amplitude ranges, and (D) number of saccades per trial per condition during the blank interval following the sequence presentation. Histograms of the horizontal and vertical eye positions peaked and were centered on the fixation at zero degrees suggesting that participants could fixate well both before and after training when predicting leftwards or rightwards oriented gratings.

Data analysis

Behavioral data analysis

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

fMRI data analysis

Pre-processing: neuroimaging data was analyzed using Brain Voyager QX (Brain Innovation, Maastricht, Netherlands). Pre-processing of functional data included slice scan time correction, three-dimensional motion correction, linear trend removal, temporal high-pass filtering (3 cycles) and spatial smoothing using a 3D Gaussian kernel of 5mm full width at half maximum (FWHM). Blocks with head motion larger than 3 mm of translation or 1º of

rotation or sharp motion above 1 mm were excluded from the analysis. The functional images were aligned to anatomical data and the complete data were transformed into Talairach space. For each observer, the functional imaging data between the two sessions were co-aligned, registering all the volumes for each observer to the first functional volume of the first run and session. This procedure ensured a cautious registration across sessions.

Whole-brain General Linear Model (GLM): The BOLD responses to structured and random before and after training sequences were modelled using a general linear model (GLM). We constructed a multiple regression design matrix that included the two stimulus conditions (structured vs. random sequences) for each of the two scanning sessions (pre- and post-training) as regressors. To remove residual motion artifacts, the six zero-centered head movement parameters were also included as regressors. We used a canonical hemodynamic response function (HRF) with a 30 s (i.e. based on the block duration) boxcar function to generate regressors. Serial correlations were corrected using a second order autoregressive model AR(2). The resulting parameter estimates (β) were used in a voxel-wise mixed design ANOVA: 2 (*session: pre- vs. post-training*) x 2 (*sequence: structured vs. random*) x 2 (*group: MCI vs. controls)*. Statistical maps were cluster threshold corrected ($p < 0.05$).

Percent Signal Change analysis (PSC) analysis: Using the whole-brain GLM analysis, we identified regions of interest (ROI) that showed significant session x sequence x group, or session x sequence interactions ($p < .05$, cluster threshold corrected). We then conducted a complementary percent signal change (PSC) analysis using a split-half data cross-validation procedure. To avoid circularity, we defined ROIs based on half of the data (odd or even runs) and calculated percent signal change (PSC) using the rest of the data. For each ROI, we calculated PSC by subtracting fMRI responses to random sequences from fMRI responses to structured sequences and dividing by the average fMRI response to random sequences. In particular, for each participant we first selected all the odd runs and conducted a GLM

analysis to identify voxels that showed learning-dependent changes within the ROIs derived from the whole brain GLM analysis. We then repeated this procedure using the data from even runs. Although the voxels selected from these two halves of the ROI-based analysis may differ, they are restricted within the same anatomical ROIs as defined by the group voxel-wise GLM analysis. Both cross-validations used independent data sets to identify voxels of interest and extract PSC signals. Thus, performing this procedure twice allowed us to cross-validate the data with at least two independent data sets resulting in more conservative and robust PSC estimates. These cross-validations resulted in similar patterns of results (see Supplementary information); therefore we averaged the PSC signals that corresponded to the same anatomical ROIs (as determined by the group voxel-wise GLM analysis) but extracted based on two independent voxel selections, avoiding circularity.

Normalizing fMRI responses by vascular reactivity: To control for possible vascular differences between the MCI patients and controls, we normalized the PSC in each ROI by a vascular reactivity measure calculated based on fMRI responses to a breath-hold task. Both MCI patients (n = 8) and controls (n = 9) carried out a breath-hold task that has been shown to cause vascular dilation induced by hypercapnia (Handwerker, Gazzaley, Inglis, & D'Esposito, 2007). In this task, the participants were instructed to hold their breath whenever they saw a black circle on the screen (10 s) or breath normally when the white circle was presented (10 s each). For each ROI, the difference in fMRI responses between the breath-hold and the normal breathing periods was used as a scaling factor to normalize the PSC for structured and random sequences.

Results

Behavioural results

We presented patients and age-matched controls with a sequence of leftwards and rightwards oriented gratings (Figure 1) and asked them to predict the next grating in the sequence. We trained participants on this prediction task without feedback for up to 5 training sessions. To discourage explicit memorization of individual item positions in the sequence, we presented all stimuli at the same rate and in a continuous stream. Further, the first item in the sequence was randomized, the last three items were the same across sequences, the position of the test stimulus was randomized across trials and for half of the trials the incorrect test stimulus was presented.

Analysis of the behavioral results showed that both MCI patients and controls improved during training (Figure S1A). Conducting this analysis with data from all participants (n=18) or with data from only 14/18 participants (n =14: MCI= 8; controls=6) who completed all five training sessions (Figure S1B) showed similar results. To compare performance between MCI patients and controls before and after training during scanning (Figure 2A), we conducted a 3-way mixed design ANOVA (*session: pre- vs. post-test*; *sequence: structured vs. random*; *group: MCI vs. controls*) including data from all participants. Participants' performance improved significantly after training (main effect of session: $F_{(1,16)} = 8.07$, $p = .012$). However, there was no significant difference between patients and controls (main effect of group: $F_{(1,16)} = 2.73$, $p = .118$), nor a significant interaction between *session, sequence* and *group* ($F_{(1,16)} = 0.094$, $p = .763$) nor between *session* and *group* ($F_{(1,16)} = 0.51$, $p = .485$). Further, we observed a main effect of *sequence* ($F_{(1,16)} = 9.47$, $p = .007$) and a significant *session* X *sequence* interaction ($F_{(1,16)} = 5.09$, $p = .038$), indicating better performance and higher improvement after training for structured than random sequences.

Similar to this analysis, a 2 (*group: MCI vs. Control)* x 6 (*session: pre, training 1,2,3,4, post-training scanner*) mixed design ANOVA including training and test sessions (Figure S1A) showed a significant effect of session ($F_{(5,80)}$ = 6.01, p < .001), but no significant interaction between session and group ($F_{(5,80)}$ =.183, p = .968). Within-subjects contrasts revealed a significant linear effect of session ($F_{(1,16)}$ = 10.5, p = .001), but no significant interaction with group ($F_{(1,16)}$ = 0.1, p = .922), suggesting that performance improved across training for both patients and controls. Interestingly, this analysis showed a marginally significant effect for group ($F_{(1,16)}$ = 3.47, p = .081), consistent with a trend of better performance for controls than patients (Figure 2A. Figure S1A). When considering performance for structured sequences alone, we observed a significant effect of group ($F_{(1,16)}$ = 6.75, p = .019) that appeared to be driven by marginally significant differences in performance between groups before (t(16) = -1.84, p = .083) but not after (t(16) = -1.68, p = .111) training.

Finally, to quantify improvement for the trained structured sequences we computed an index subtracting pre-training from post-training performance (Figure 2B). A 2 (*sequence: structured vs. random*) x 2 (*group:* MCI *vs.* controls) mixed design ANOVA showed a significant effect for *sequence* ($F_{(1,16)}$ = 5.09, p =.038), but no significant effect of group ($F_{(1,16)}$ = 0.51, p =.485) nor a significant interaction between *sequence* and *group* ($F_{(1,16)}$ = 0.09, p =.763). Thus, despite differences between groups in task performance before training, this analysis indicates similar training-dependent performance improvement for MCI patients and controls that is specific to the trained structured rather than random sequences.

FIGURE 2 ABOUT HERE

fMRI results

To identify brain regions that show learning-dependent changes in the prediction task, we scanned participants before and after training and compared BOLD responses for structured vs. random sequences across sessions. To assess differences in learning-dependent fMRI changes between patients and controls, we followed two complementary analyses procedures. First, we identified brain regions that showed significant learning-dependent changes in fMRI signals, using a whole brain voxel-wise GLM analysis (RFX group analysis). We then used these regions as anatomical ROIs of interest for further analyses of the fMRI responses (PSC: percent signal change) to understand differences between conditions (structured vs. random sequences) and groups (patients vs. controls) before and after raining. The whole brain GLM analysis (RFX, cluster-threshold corrected at $p<0.05$) revealed a network of cortical and subcortical brain regions (Figures 3, 4) that showed differences in BOLD responses between structured and random sequences after training. We focus on the 3-way interaction of *group* x *session* x *sequence* (Table 1) to identify regions showing differences between patients and controls in training-induced improvement. We further test for regions showing a significant two-way *session* x *sequence* interaction to test for any additional brain regions showing training-induced improvement across participants (Table 2). The results from main effects and other interactions are shown in additional supplementary tables (Table S1 - S5). To interpret brain activations resulting from these interactions, for each participant and each of the regions identified from the whole brain GLM analysis we extracted fMRI responses (PSC: percent signal change) for structured compared to random sequences before and after training (see methods for details). To avoid circularity in this region of interest analysis, we used a split-half cross-validation procedure. That is, we used only half the data (first the odd, then the even runs) to identify regions of interest in each participant and the rest of the data to extract PSC. This iterative procedure of selecting voxels within an ROI and evaluating their

signals provides a conservative way of cross-validating our results and allows us to identify key brain regions that show robust learning-dependent fMRI changes between groups. Below, we focus on ROIs that had volume higher than 300 mm^3 and showed a significant ($p < .05$) interaction (session x sequence x group or session x sequence) for the split-half data procedure.

To compare directly differences in learning-dependent fMRI changes in patients and controls, we tested for regions that showed a significant interaction between group x session x sequence. The whole brain voxel-wise GLM analysis showed a network of frontal, middle temporal, subcortical (parahippocampus, thalamus, basal ganglia), and cerebellar regions (Figure 3, Table 1). Further, the complementary PSC analysis using a split-half cross-validation procedure showed significantly increased fMRI responses for structured sequences after training for MCI patients compared to age-matched controls (Figure 3 for mean PSC data across cross-validations; Figure S2 for PSC data per cross-validation) in a network of regions in frontal (SFG and MiFG), subcortical (THM, PTM), and cerebellar regions (CL, CT, and CUL). This was confirmed by a mixed design 2 (session: pre- vs. post-training) x 2 (group: MCI vs. Control) ANOVA that showed significant interactions between session and group in frontal ($F_{(1,16)} = 6.07$, $p = .025$), subcortical (THM and PTM: $F_{(1,16)} = 13.88$, $p = .002$), and cerebellar ($F_{(1,16)} = 18.38$, $p = .001$) regions. As shown in Figure 3, these results suggest enhanced fMRI responses for structured sequences after training (frontal: $F_{(1,8)} = 6.49$, $p = 0.032$; subcortical: $F_{(1,8)} = 17.14$, $p = 0.003$; cerebellar: $F_{(1,8)} = 22.62$, $p = 0.001$) for MCI patients, but not for age-matched controls (frontal: $F_{(1,8)} = 0.24$, $p = 0.628$; subcortical: $F_{(1,8)} = 0.18$, $p = 0.683$; cerebellar: $F_{(1,8)} = 1.30$, $p = 0.287$). Further, post-hoc comparisons showed significant differences in fMRI responses between participant groups (i.e. higher PSC for structured sequences in patients that controls) after training in these regions: SFG ($t_{(1,16)} = 2.75$, $p = 0.014$), MiFG ($t_{(1,16)} = 2.55$, $p = 0.021$), THM ($t_{(1,16)} = 3.13$, $p = 0.004$), PTM ($t_{(1,16)}$

= 3.45, p = 0.003), CUL ($t_{(1,16)}$ = 3.96, p = 0.001), CL ($t_{(1,16)}$ = 2.84, p = 0.012), CT ($t_{(1,16)}$ = 2.35, p = 0.032). However, no significant differences were observed in fMRI responses in these regions (P>0.05) between patients and controls before training, suggesting that differences in fMRI responses between groups could not be attributed to the marginally –but not statistically significant- higher performance for controls before training.

FIGURE 3 ABOUT HERE

We then tested for any additional brain regions showing training-induced improvement across patients and controls (Table 2). The whole brain voxel-wise GLM analysis showed a more extended network of frontal, medial posterior, temporal, subcortical (hippocampus, thalamus, basal ganglia, claustrum), and cerebellar regions (Figure 4, Table 2). We further conducted PSC analysis in these regions, using a mixed design 2 (session: pre- vs. post-training) x 2 (group: MCI vs. Control) ANOVA for each region. Our results (Figure 4 for mean PSC data across cross-validations; Figure S3 for PSC data per cross-validation) showed a significant effect of *session* in frontal ($F_{(1,16)}$ = 10.59, p =.005), posterior ($F_{(1,16)}$ = 8.76, p =.009), temporal ($F_{(1,16)}$ = 7.35, p =.015), subcortical ($F_{(1,16)}$ = 7.34, p =.015), and cerebellar ($F_{(1,16)}$ = 11.11, p =.004), but no significant main effects of group (p > .05). These results are consistent with training-induced improvement across participants. We then tested for interactions between session and group; these were not significant for frontal, posterior, and temporal regions (p > 0.05). However, we observed a significant interaction between *session* and *group* in cerebellar regions ($F_{(1,16)}$ = 8.59, p =0.010), confirming enhanced fMRI responses in patients for structured sequences after training. We also observed a significant three-way interaction between session x ROI (HC and CDT) x group ($F_{(1,16)}$ = 5.96, p =.027) for subcortical regions. Pairwise comparisons following this interaction showed significantly increased fMRI responses for structured sequences after training in hippocampus for controls ($t_{(1,8)}$ = 5.62, p <.001), but not patients($t_{(1,8)}$ = -0.53, p =.612) . The same pairwise contrasts

did not show significant learning-dependent changes in caudate for patients ($t_{(1,8)}$ = 1.82, p =.106), nor controls ($t_{(1,8)}$ = 1.84, p =.102),

FIGURE 4 ABOUT HERE

Taken together these analyses reveal a network of frontal, subcortical and cerebellar regions that shows learning-dependent changes in fMRI responses primarily for MCI patients. Interestingly, we observed stronger learning-dependent changes for controls in the hippocampus, although effects in this anatomically smaller region appeared weaker. In particular, a 2-way GLM interaction between session and group (Table 2) showed significant activation in subgyral hippocampus. Further, PSC analysis in this region showed stronger learning-dependent changes for controls rather than patients. These results suggest that when hippocampal processing is disrupted -as in the case of MCI- an alternate network including cortical and subcortical regions may be employed to support learning of predictive structures.

Further, we tested whether these learning-dependent changes in fMRI responses for structured sequences could be explained by differences in the fMRI response to random sequences. Analysing PSC data from fixation baseline for random sequences did not show any significant interactions between *session* and *group* in regions of interest showing differences in learning-dependent fMRI changes between MCI patients and controls: frontal ($F_{(1,16)}$ = 2.05, p =.171), subcortical ($F_{(1,16)}$ = 1.70, p =.211), or cerebellar ($F_{(1,16)}$ = 1.06, p =.318) areas. Further, there was no significant effect for *session* (frontal: $F_{(1,16)}$ = 0.21, p =.650; subcortical: $F_{(1,16)}$ = 2.95, p =.105; cerebellar: $F_{(1,16)}$ = 0.05, p =.826) nor *group* (frontal: $F_{(1,16)}$ = 0.13, p =.726; subcortical: $F_{(1,16)}$ = 0.44, p =.515; cerebellar: $F_{(1,16)}$ = 2.42, p =.139) in the PSC in response to random sequences. Therefore, the learning-dependent changes we observed for fMRI responses to structured sequences cannot be explained by

differences in the baseline (i.e. fMRI responses to random sequences) before vs. after training.

Finally, we asked whether the learning-dependent changes we observed were linked to behavioral improvement and may therefore reflect compensatory activity. Correlations between performance and fMRI changes (i.e. difference before vs. after training) did not reach significance. Further, we tested correlations of ACE-III scores and learning-dependent fMRI changes (post-training minus pre-training PSC). The battery of tests included in ACE-III measures general cognitive abilities (including memory, executive function, visuospatial skills, attention and language) and can be therefore used as an independent measure of cognitive capacity. This analysis showed a significant positive correlation in putamen for patients ($r = .776$, $p = 0.014$), but not controls ($r = -.196$, $p = 0.613$). Comparison of these two independent correlations (Fisher Z-test) showed a significant difference between patients and controls ($Z = -2.14$, $p = 0.0325$). This analysis shows that enhanced activity in putamen after training may relate to higher cognitive capacity, suggesting that patients with high cognitive capacity may recruit putamen to improve in learning predictive structures despite hippocampal dysfunction. However, given the small number of participants in each group, future work will need to test this hypothesis further with larger numbers of participants.

Control Analyses

We conducted the following additional analyses and experiments to control for possible alternative explanations of the results.

First, we asked whether the differences in fMRI responses between structured and random sequences were due to the participants attending more to the structured sequences after training. Debriefing the participants after training suggests that this is unlikely, as the participants were not aware that some of the sequences were structured and some others

random. Further, comparing response times to structured and random sequences in the pre- and post-training session between the two groups (3 way mixed design ANOVA: *session* X *sequence* X *group*) showed decreased response times after training (main effect of *session*: $F_{(1,8)}$ = 10.29, p =.005), but no significant differences between structured and random sequences (main effect of *sequence* $F_{(1,8)}$ = 0.043, p =.838), suggesting that participants engaged with the task when both structured and random sequences were presented. Importantly, there was no significant interaction between *session and group* $F_{(1,8)}$ = 1.40, p =.254 nor a three-way interaction (*session* X *sequence* X *group*, $F_{(1,8)}$ = 0.015, p =.904), suggesting that our results couldn't be simply due to differences in processing speed or sensorimotor performance.

Second, is it possible that the differences in activation patterns observed between patients and controls originated from differences in vascular reactivity rather than differences in underlying neuronal activity (D'Esposito, Deouell, & Gazzaley, 2003; D'Esposito, Zarahn, Aguirre, & Rypma, 1999; Restom, Bangen, Bondi, Perthen, & Liu, 2007) (Hamzei, Knab, Weiller, & Rother, 2003). Although, no patients with vascular-related disease were included in the study, to control for this possibility, we acquired fMRI data in both patients and controls during a breath-holding task (Handwerker et al., 2007). The BOLD signal change induced by the hypercapnic challenge of this task was used as an estimate of the vascular reactivity in every voxel. For each participant, we used the BOLD response amplitude to the breath-holding task (i.e. difference in BOLD signal during ten-second breath-holding vs. normal breathing periods) to normalize the stimulus evoked BOLD signal, as previously described (Handwerker et al., 2007). For each region of interest shown in Figures 3 and 4, we divided the percent BOLD signal evoked by the experimental task by the percent BOLD evoked by the hypercapnic breath-holding task. These normalized fMRI responses (Figure 5) showed a similar pattern of results as in Figures 3 and 4, suggesting that the differences in

activation patterns that we observed between patients and controls could not be simply explained by differences in vascular reactivity between participant groups. Further, lack of differences in brain patterns between groups could be due to insufficient sensitivity of fMRI recordings in these regions. Analysis of functional signal-to-noise ratio SNR (Figure S4) in these regions demonstrates that we recorded fMRI signals with similar sensitivity in patients and controls, allowing us to compare between brain regions and participant groups.

FIGURE 4 ABOUT HERE

Third, we tested whether differences in brain activations between participant groups were due to baseline differences in performance before training. Our behavioural results showed similar behavioural improvement for patients and controls (as indicated by lack of a significant interaction between *group* and *session)* allowing direct comparison of fMRI activity patterns between participant groups. Further, correlating pre-training behavioral performance with training-dependent fMRI changes (PSC after-PSC before training) did not show any significant correlations in any of the ROIs considered in this analysis (Table S6). Thus, these results suggest that differences in brain activation patterns between participant groups couldn't be simply attributed to differences in task performance before training.

Finally, analysis of eye-movement recordings did not show any significant differences between structured and random sequences for both patients and controls (Figure S5). In particular, we did not observe any effect of sequence for the mean eye position ($F_{(1,4)} = 0.043$, $p = .846$), mean saccade amplitude ($F_{(1,4)} = 0.312$, $p = .606$), or number of saccades per trial per condition ($F_{(1,4)} = 0.477$, $p = .528$). There was also no significant effect for group on the mean eye position ($F_{(1,4)} = 2.67$, $p = .178$), mean saccade amplitude ($F_{(1,4)} = 0.464$, $p = .533$), or number of saccades per trial per condition ($F_{(1,4)} = 0.883$, $p = .401$). Finally, there were no significant interactions between *sequence* and *group* ($p > 0.05$). These results suggest that it is unlikely that differences in eye movements between sequences (e.g. structured sequences

may be more interesting than random sequences and incur more eye movements) or participant groups (e.g. patients getting more easily tired or distracted during scanning could make more eye movements) could explain differences in fMRI signals between sequences or participant groups.

Discussion

Our results demonstrate that MCI-AD patients benefit from exposure to temporal sequences and acquire knowledge that allows them to predict upcoming events. Both patients and controls showed similar improvement after training in the prediction task allowing us to compare directly brain activity between the two groups. Our fMRI results provide evidence that a frontal-subcortical-cerebellar circuit- that is known to be involved in implicit learning of temporal structures- mediates the ability of MCI-AD patients to accumulate information about temporal regularities through repeated exposure and predict future events. Our findings advance our understanding of cognitive function in MCI-AD in three main respects.

First, our study is the first to test 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 & Bullemer, 1987); for review see (Schwarb & 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 pre-determined 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, Yokomizo, & Bottino, 2012). In contrast, using an explicit prediction test, we demonstrate that predictions related to identification of objects are facilitated by implicit knowledge of temporal structures. Our findings are consistent with previous studies suggesting that learning of regularities may occur implicitly in a range of tasks: visuomotor sequence learning (Nissen & Bullemer, 1987), artificial grammar learning (Reber, 1967), probabilistic category learning (Knowlton, Squire, & Gluck, 1994), and contextual cue learning (Chun & 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 (leftward vs. rightward 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, suggesting implicit knowledge of the trained temporal structures.

Second, our study provides novel evidence for a brain circuit –alternate to hippocampus- that may support learning of predictive structures despite hippocampal dysfunction in MCI. Previous behavioural studies showing preserved implicit learning in MCI patients (Negash et al., 2007; Nemeth et al., 2013), have not investigated the brain circuits that may support this ability in MCI. Our results suggest that predictive learning engages a cortico-striatal-cerebellar network that is thought to be involved in implicit learning of temporal regularities and is spared in MCI. These findings are consistent with previous work implicating subcortical areas in associative learning and temporal memory, while prefrontal circuits in rule-based behaviours and prediction of future events (Bar, 2009; Leaver, Van Lare,

Zielinski, Halpern, & Rauschecker, 2009; Pasupathy & Miller, 2005). In particular, striatal regions have been implicated mainly in implicit learning (Hazeltine, Grafton, & Ivry, 1997; Rauch et al., 1995), while the medial temporal lobe in both implicit and explicit learning (Schendan et al., 2003a; Schendan, Searl, Melrose, & Stern, 2003b). Finally, our results are consistent with recent work implicating cerebellum in sequence learning (e.g. serial reaction task) (Dirnberger, Novak, & Nasel, 2013; Tzvi, Muente, & Kraemer, 2014) and predictive processing (Kotz, Stockert, & Schwartze, 2014; Leggio & Molinari, 2014).

Third, our findings are consistent with previous reports of compensatory mechanisms against brain loss in MCI-AD. In particular, we showed increased activations in frontal, subcortical (thalamus, putamen) and cerebellar regions for trained structured sequences in MCI patients compared to controls. Such over-activation patterns have been previously observed in MCI and have been associated with compensatory mechanisms against grey matter loss; for example in the hippocampus for memory (Dickerson, Salat, Greve, Albert, et al., 2005; Dickerson, Salat, Greve, Chua, et al., 2005; Hämäläinen et al., 2007) and prefrontal cortex for verbal learning (Clément & Belleville, 2010). Interestingly, previous studies have shown volume reduction in thalamus (Chen & Herskovits, 2006; Karas et al., 2004; Pennanen et al., 2005) and abnormal resting state connectivity with other brain areas (Cai et al., 2015; Wang et al., 2012; Zhou et al., 2013) for MCI patients. Further, volume reductions in thalamus and putamen, have been shown to predict cognitive decline in ageing and conversion to Alzheimer's disease (de Jong et al., 2008). Thus, it is possible that increased activations in these regions for predictive learning suggest compensatory mechanisms against grey matter loss in MCI-AD.

Several questions remain open and require further investigation. First, MCI patients have been shown to learn simple stimulus-response associations but are impaired in generalizing learning to new contexts, consistent with hippocampal dysfunction (Nagy et al., 2007). Our

study involved learning of temporal associations but did not manipulate learning context. Successful performance in the prediction task required that the participants learned temporal order statistics across the items presented in the sequences, as the frequency of occurrence was matched for the two grating orientations. 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. Future work using probabilistic sequences and manipulating the context of learning will be important for testing generalization of learning in novel situations. MCI-AD patients relying on implicit learning that is thought to be context-dependent may show impairments in generalizing to new contexts (Jimenez, Lupianez, & Vaquero, 2009). Second, 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 in MCI-AD patients. 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 brain circuits involved in predictive learning may play a compensatory role against conversion to AD.

Figures

Figure 1: Stimuli and Trial design. The trial design: a sequence of 8 gratings was presented twice. A stimulus cue followed by a test grating was presented at a random position during the second repeat of the sequence. Following the response to the test stimulus, the remaining gratings of the sequence were presented. A black 'X' indicated the end of the trial.

Figure 2: Behavioral performance. **A.** Average performance (proportion correct) across participants (MCI patients vs. controls) for the structured sequences during pre- and post-training scanning sessions. **B.** Performance improvement for MCI patients and controls for structured and random sequences calculated by subtracting the average pre-training from the average post-training performance during scanning.

Figure 3: fMRI data: three-way interaction. A. GLM maps for the three-way interaction session x sequence x group, at $p < 0.05$ (cluster threshold corrected). **B.** PSC for structured sequences before (blue) and after (red) training in MCI patients and controls. Data is shown for ROIs that had more than 300 mm^3 and showed a significant ($p < .05$) interaction between session (pre- vs. post-training), sequences (structured vs. random) and group (MCI vs. controls) for the split-half data procedure. SFG: superior frontal gyrus; MiFG: middle frontal gyrus; THM: thalamus; PTM: putamen; CUL: culmen; CL: cerebellar lingual; CT: cerebellar tonsil.

Figure 4: fMRI data: two-way interaction. A. GLM maps for the two-way interaction session x sequence, at $p < 0.05$ (cluster threshold corrected). **B.** PSC for structured sequences before (blue) and after (red) training in MCI patients and controls. Data is shown for ROIs that had more than 300 mm^3 and showed a significant ($p < .05$) interaction between session (pre- vs. post-training) and sequence (structured vs. random) for the split-half data procedure. MiFG: middle frontal gyrus; IFG: inferior frontal gyrus; MFG: medial frontal gyrus; PrG:

precentral gyrus; PsG: postcentral gyrus; STG: superior temporal gyrus; PCC: posterior cingulate cortex; LG: lingual gyrus; DEC: declive; CUL: culmen; HC: hippocampus; CDT: caudate.

Figure 5: fMRI responses normalized by vascular reactivity. For each ROI, we used differences in the fMRI response between breath-holding and the normal breathing periods to normalize the PSC for structured and random sequences. Data is shown for ROIs that had more than 300 mm^3 and showed A. a significant ($p < 0.05$) interaction between session (pre- vs. post-training) sequence (structured vs. random) and group (MCI vs. Controls) for the split-half data procedure. B. a significant ($p < 0.05$) interaction between session (pre- vs. post-training) and sequence (structured vs. random) for the split-half data procedure;

Tables

Table 1. Brain regions showing significant interaction between *session* X *sequence X group* ($p < 0.05$, cluster corrected).

ROI	H.	Volume (mm^3)	X	Y	Z	F	P
Frontal							
Superior Frontal Gyrus (SFG)	L	1569	-4	1	66	23.66843	0.000172
Superior Frontal Gyrus (SFG)	R	1967	20	-1	66	15.66589	0.001127
Medial Frontal Gyrus (MFG)	L	1231	-16	-11	48	18.94065	0.000494
Medial Frontal Gyrus (MFG)	R	946	1	-8	67	11.47566	0.003757
Middle Frontal Gyrus (MiFG)	L	126	-18	19	56	8.325398	0.010763
Middle Frontal Gyrus (MiFG)	R	840	53	25	24	13.03662	0.002346
Precentral Gyrus (PrG)	L	343	29	-17	63	16.55506	0.000893
Precentral Gyrus (PrG)	R	337	29	-17	63	16.55506	0.000893
Cingulate							
Cingulate Gyrus (CG)	L	324	-1	-2	24	10.04878	0.005941
Cingulate Gyrus (CG)	R	325	8	-2	24	15.91236	0.001056
Temporal							
Middle Temporal_Gyrus (MTG)	L	236	-49	-65	21	9.845961	0.006355
Middle Temporal_Gyrus (MTG)	R	275	46	-70	21	9.282434	0.007691
Subcortical							
Thalamus (THM)	L	533	-1	-5	0	16.59416	0.000884
Thalamus (THM)	R	447	1	-8	3	10.68645	0.004824
Caudate (CDT)	L	159	-13	-11	27	9.041909	0.008357
Caudate (CDT)	R	259	9	-2	23	9.970459	0.006097
Putamen (PTM)	L	297	-19	7	-3	7.661206	0.01372
Putamen_(PTM)	R	98	30	-11	-9	16.01083	0.001029
Parahippocampal Gyrus (PHG)	R	848	29	-11	-12	39.11372	0.000012
Cerebellar							
Cerebellar Lingual (CL)	L	316	-4	-44	-18	14.68998	0.001468
Cerebellar Lingual (CL)	R	255	0	-44	-15	13.57174	0.002009
Cerebellar Tonsil (CT)	L	304	-4	-56	-33	8.005397	0.012085
Cerebellar Tonsil (CT)	R	33	2	-50	-33	5.62813	0.030551
Culmen (CUL)	L	563	-4	-43	-18	13.75714	0.001906
Culmen (CUL)	R	872	11	-38	-12	14.0551	0.001751

Table 2. Brain regions showing significant interaction between *session and sequence* ($p < 0.05$, cluster corrected).

	H.	Size	X	Y	Z	F	P
Frontal							
Insula (INS)	L	1526	-46	4	6	19.71843	0.000411
Insula (INS)	R	595	32	-23	21	14.03154	0.001763
Superior Frontal Gyrus (SFG)	L	1167	-4	-2	67	14.21171	0.001676
Superior Frontal Gyrus (SFG)	R	112	1	1	64	7.770197	0.013176
Inferior Frontal Gyrus (IFG)	L	1088	-49	7	27	17.43258	0.000715
Medial Frontal Gyrus (MFG)	L	287	-1	1	63	12.0809	0.003119
Medial Frontal Gyrus (MFG)	R	187	2	-26	60	8.953101	0.008619
Middle Frontal Gyrus (MiFG)	L	520	-42	-2	48	9.789592	0.006476
Precentral Gyrus (PrG)	L	2421	-46	3	6	18.22149	0.000588
Precentral Gyrus (PrG)	R	142	56	-5	12	8.028649	0.011983
Postcentral Gyrus (PsG)	L	758	-58	-29	21	15.74848	0.001103
Postcentral Gyrus (PsG)	R	302	59	-17	18	15.10464	0.001311
Medial Posterior							
Posterior Cingulate (PCC)	L	1426	-10	-47	12	10.94162	0.004446
Posterior Cingulate (PCC)	R	1193	5	-53	12	12.18002	0.003027
Lingual Gyrus (LG)	L	1133	-11	-52	0	12.70769	0.002585
Lingual Gyrus (LG)	R	270	29	-57	-2	7.364342	0.015334
Precuneus (PREC)	L	265	-1	-62	54	9.137805	0.008084
Precuneus (PREC)	R	871	11	-53	51	8.897297	0.008789
Cingulate Gyrus (CG)	L	679	-1	-5	24	12.57344	0.00269
Cingulate Gyrus (CG)	R	716	5	1	24	21.73441	0.00026
Parietal-Occipital							
Fusiform Gyrus (FsG)	L	543	-37	-74	-12	13.54358	0.002025
Fusiform Gyrus (FsG)	R	54	29	-56	-5	9.075971	0.008259
Inferior Parietal Lobule (IPL)	L	482	-58	-30	21	14.10169	0.001729
Temporal							
Superior Temporal Gyrus (STG)	L	1405	-37	-53	12	13.36493	0.002132
Superior Temporal Gyrus (STG)	R	14	51	-38	6	5.095407	0.038323
Transverse Temporal Gyrus (TTG)	L	289	-58	-14	12	9.069391	0.008278
Transverse Temporal Gyrus (TTG)	R	188	59	-14	15	10.13854	0.005767
Subcortical							
Parahippocampal Gyrus (PHG)	L	1384	-31	-39	-3	18.22406	0.000587
Parahippocampal Gyrus (PHG)	R	2117	32	-11	-15	25.65191	0.000115
Sub-gyral Hippocampus (HC)	R	551	32	-42	0	11.55167	0.00367
Caudate (CDT)	L	328	-34	-42	1	14.53081	0.001534
Caudate (CDT)	R	2923	19	1	21	19.68948	0.000414
Putamen (PTM)	L	2970	-25	10	0	26.45068	0.000098
Putamen (PTM)	R	104	21	1	20	14.25854	0.001654
Thalamus (THM)	L	783	-4	-20	18	13.52562	0.002036
Thalamus (THM)	R	1265	5	-23	18	17.53045	0.000697
Claustrum (CLAU)	L	318	-22	24	9	12.39722	0.002835
Claustrum (CLAU)	R	88	30	-17	20	7.629029	0.013885
Cerebellar							
Culmen (CUL)	L	3007	-7	-50	0	22.51225	0.00022
Culmen (CUL)	R	2257	11	-59	-9	14.88503	0.001391
Declive (DEC)	L	1031	-40	-71	-15	12.71706	0.002577
Declive (DEC)	R	1019	26	-78	-18	17.53406	0.000697

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 workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers & Dementia, 7*(3), 270-279. doi: 10.1016/j.jalz.2011.03.008

Aslin, R. N., & Newport, E. L. (2012). Statistical Learning: From Acquiring Specific Items to Forming General Rules. *Current Directions in Psychological Science, 21*(3), 170-176. doi: 10.1177/0963721412436806

Baker, R., Dexter, M., Hardwicke, T. E., 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, G. L., Albert, M. S., Speck, C. L., Jones, L. R., Stark, C. E., . . . Gallagher, M. (2012). Reduction of hippocampal hyperactivity improves cognition in amnestic mild cognitive impairment. *Neuron, 74*(3), 467-474. doi: 10.1016/j.neuron.2012.03.023

Bar, M. (2009). The proactive brain: memory for predictions. *Philos Trans R Soc Lond B Biol Sci, 364*(1521), 1235-1243. doi: 10.1098/rstb.2008.0310

Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology-Human Perception and Performance, 33*(4), 798-815. doi: 10.1037/0096-1523.33.4.798

Brady, T. F., & Oliva, A. (2008). Statistical learning using real-world scenes - Extracting categorical regularities without conscious intent. *Psychological Science, 19*(7), 678-685. doi: 10.1111/j.1467-9280.2008.02142.x

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat Vis, 10*(4), 433-436.

Cai, S., Huang, L., Zou, J., Jing, L., Zhai, B., Ji, G., . . . for the Alzheimer's Disease Neuroimaging, I. (2015). Changes in Thalamic Connectivity in the Early and Late Stages of Amnestic Mild Cognitive Impairment: A Resting-State Functional Magnetic Resonance Study from ADNI. *PLoS One, 10*(2), e0115573. doi: 10.1371/journal.pone.0115573

Celone, K. A., Calhoun, V. D., Dickerson, B. C., Atri, A., Chua, E. F., Miller, S. L., . . . Sperling, R. A. (2006). Alterations in memory networks in mild cognitive impairment and Alzheimer's disease: An independent component analysis. *Journal of Neuroscience, 26*(40), 10222-10231. doi: 10.1523/jneurosci.2250-06.2006

Chen, R., & Herskovits, E. H. (2006). Network analysis of mild cognitive impairment. *Neuroimage, 29*(4), 1252-1259. doi: http://dx.doi.org/10.1016/j.neuroimage.2005.08.020

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*(5), 170-178. doi: 10.1016/s1364-6613(00)01476-5

Chun, M. M., & Jiang, Y. H. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28-71. doi: 10.1006/cogp.1998.0681

Clément, F., & Belleville, S. (2010). Compensation and Disease Severity on the Memory-Related Activations in Mild Cognitive Impairment. *Biol Psychiatry, 68*(10), 894-902. doi: http://dx.doi.org/10.1016/j.biopsych.2010.02.004

Curran, T. (1997). Effects of aging on implicit sequence learning: Accounting for sequence structure and explicit knowledge. *Psychology Research, 60*, 24 - 41.

D'Esposito, M., Deouell, L. Y., & Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. *Nat Rev Neurosci, 4*(11), 863-872.

D'Esposito, M., Zarahn, E., Aguirre, G. K., & Rypma, B. (1999). The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. *Neuroimage, 10*(1), 6-14.

de Jong, L. W., van der Hiele, K., Veer, I. M., Houwing, J. J., Westendorp, R. G. J., Bollen, E. L. E. M., . . . van der Grond, J. (2008). Strongly reduced volumes of putamen and thalamus in Alzheimers disease: an MRI study. *Brain, 131*, 3277-3285. doi: Doi 10.1093/Brain/Awn278

Dickerson, B. C., Salat, D. H., Bates, J. F., Atiya, M., Killiany, R. J., Greve, D. N., . . . Sperling, R. A. (2004). Medial temporal lobe function and structure in mild cognitive impairment. *Ann Neurol, 56*(1), 27-35. doi: 10.1002/ana.20163

Dickerson, B. C., Salat, D. H., Greve, D. N., Albert, M. S., Blacker, D., & Sperling, R. A. (2005). Memory-related medial temporal lobe activation in mild cognitive impairment prior to dementia: An fMRI study. *Neurology, 64*(6), A227-A227.

Dickerson, B. C., Salat, D. H., Greve, D. N., Chua, E. F., Rand-Giovannetti, E., Rentz, D. M., . . . Sperling, R. A. (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

Dirnberger, G., Novak, J., & Nasel, C. (2013). Perceptual Sequence Learning Is More Severely Impaired than Motor Sequence Learning in Patients with Chronic Cerebellar Stroke. *Journal of Cognitive Neuroscience, 25*(12), 2207-2215. doi: 10.1162/jocn_a_00444

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature, 427*(6972), 311-312.

Eichenbaum, H. (2013). Memory on time. *Trends Cogn Sci, 17*(2), 81-88. doi: 10.1016/j.tics.2012.12.007

Fiser, J., & Aslin, R. N. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology-Learning Memory and Cognition, 28*(3), 458-467. doi: 10.1037//0278-7393.28.3.458

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proc Natl Acad Sci U S A, 103*(31), 11778-11783. doi: 10.1073/pnas.0602659103

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

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585–612.

Hämäläinen, A., Pihlajamäki, M., Tanila, H., Hänninen, T., Niskanen, E., Tervo, S., . . . Soininen, H. (2007). Increased fMRI responses during encoding in mild cognitive impairment. *Neurobiol Aging, 28*(12), 1889-1903. doi: http://dx.doi.org/10.1016/j.neurobiolaging.2006.08.008

Hamzei, F., Knab, R., Weiller, C., & Rother, J. (2003). The influence of extra- and intracranial artery disease on the BOLD signal in FMRI. *Neuroimage, 20*(2), 1393-1399. doi: 10.1016/S1053-8119(03)00384-7 S1053811903003847 [pii]

Handwerker, D. A., Gazzaley, A., Inglis, B. A., & D'Esposito, M. (2007). Reducing vascular variability of fMRI data across aging populations using a breathholding task. *Hum Brain Mapp, 28*(9), 846-859. doi: 10.1002/hbm.20307

Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristics determine the locus of motor-sequence encoding. A PET study. *Brain, 120 (Pt 1)*, 123-140.

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *J Exp Psychol Gen, 142*(4), 1211-1230. doi: 10.1037/a0033621

Hsieh, L. T., Gruber, M. J., Jenkins, L. J., & 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

Ingvar, D. H. (1985). "Memory of the future": an essay on the temporal organization of conscious awareness. *Hum Neurobiol, 4*(3), 127-136.

Jimenez, L., Lupianez, J., & Vaquero, J. M. M. (2009). Sequential congruency effects in implicit sequence learning. *Consciousness and Cognition, 18*(3), 690-700. doi: 10.1016/j.concog.2009.04.006

Karas, G. B., Scheltens, P., Rombouts, S. A. R. B., Visser, P. J., van Schijndel, R. A., Fox, N. C., & Barkhof, F. (2004). Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. *Neuroimage, 23*(2), 708-716. doi: http://dx.doi.org/10.1016/j.neuroimage.2004.07.006

Kemp, C., & Tenenbaum, J. B. (2009). Structured Statistical Models of Inductive Reasoning (vol 116, pg 20, 2009). *Psychological Review, 116*(2), 461-461. doi: 10.1037/a0015514

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & memory (Cold Spring Harbor, N.Y.), 1*(2), 106-120.

Kotz, S. A., Stockert, A., & Schwartze, M. (2014). Cerebellum, temporal predictability and the updating of a mental model. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 369*(1658). doi: 10.1098/rstb.2013.0403

Kourtzi, Z. (2010). Visual learning for perceptual and categorical decisions in the human brain. *Vision Res, 50*(4), 433-440. doi: S0042-6989(09)00453-2 [pii]10.1016/j.visres.2009.09.025

Kunda, Z., & Nisbett, R. E. (1986). The psychometrics of everyday life. *Cognitive Psychology, 18*(2), 195-224. doi: 10.1016/0010-0285(86)90012-5

Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *J Neurosci, 29*(8), 2477-2485. doi: 10.1523/JNEUROSCI.4921-08.2009

Leggio, M., & Molinari, M. (2014). Cerebellar Sequencing: a Trick for Predicting the Future. *Cerebellum*.

Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). Sequential Expectations: The Role of Prediction-Based Learning in Language. *Topics in Cognitive Science, 2*(1), 138-153. doi: 10.1111/j.1756-8765.2009.01072.x

Morris, J. C., & Cummings, J. (2005). Mild cognitive impairment (MCI) represents early-stage Alzheimer's disease. *Journal of Alzheimers Disease, 7*(3), 235-239.

Nagy, H., Keri, S., Myers, C. E., Benedek, G., Shohamy, D., & Gluck, M. A. (2007). Cognitive sequence learning in Parkinson's disease and amnestic mild cognitive impairment: Dissociation between sequential and non-sequential learning of associations. *Neuropsychologia, 45*(7), 1386-1392. doi: 10.1016/j.neuropsychologia.2006.10.017

Negash, S., Petersen, L. E., Geda, Y. E., Knopman, D. S., Boeve, B. E., Smith, G. E., . . . Petersen, R. C. (2007). Effects of ApoE genotype and mild cognitive impairment on implicit learning. *Neurobiol Aging, 28*(6), 885-893. doi: DOI 10.1016/j.neurobiolaging.2006.04.004

Nemeth, D., Janacsek, K., Kiraly, K., Londe, Z., Nemeth, K., Fazekas, K., . . . Csanyi, A. (2013). Probabilistic sequence learning in mild cognitive impairment. *Frontiers in Human Neuroscience, 7*. doi: Unsp 318 Doi 10.3389/Fnhum.2013.00318

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning - evidence from performance measures. *Cognitive Psychology, 19*(1), 1-32. doi: 10.1016/0010-0285(87)90002-8

Pasupathy, A., & Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature, 433*(7028), 873-876. doi: 10.1038/nature03287

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis, 10*(4), 437-442.

Pennanen, C., Testa, C., Laakso, M. P., Hallikainen, M., Helkala, E.-L., Hänninen, T., . . . Soininen, H. (2005). A voxel based morphometry study on mild cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry, 76*(1), 11-14. doi: 10.1136/jnnp.2004.035600

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, R. C., Roberts, R. O., Knopman, D. S., Boeve, B. F., Geda, Y. E., Ivnik, R. J., . . . Jack Jr, C. R. (2009). Mild Cognitive Impairment: Ten Years Later. *Neurological Review, 66*(12).

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment - Clinical characterization and outcome. *Archives of Neurology, 56*(3), 303-308. doi: 10.1001/archneur.56.3.303

Pirogovsky, E., Holden, H. M., Jenkins, C., Peavy, G. M., Salmon, D. P., Galasko, D. R., & Gilbert, P. E. (2013). Temporal Sequence Learning in Healthy Aging and Amnestic Mild Cognitive Impairment. *Experimental Aging Research, 39*(4), 371-381. doi: 10.1080/0361073x.2013.808122

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature, 414*(6863), 546-550. doi: 10.1038/35107080

Poldrack, R. A., Prabhakaran, V., Seger, C. A., & Gabrieli, J. D. (1999). Striatal activation during acquisition of a cognitive skill. *Neuropsychology, 13*(4), 564-574.

Rauch, S. L., Savage, C. R., Brown, H. D., Curran, T., Alpert, N. M., Kendrick, A., . . . Kosslyn, S. M. (1995). A PET investigation of implicit and explicit sequence learning. *Human Brain Mapping, 3*(4), 271-286. doi: 10.1002/hbm.460030403

Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., Brown, H. D., . . . Rosen, B. R. (1997). Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Hum Brain Mapp, 5*(2), 124-132.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*(6), 855-&. doi: 10.1016/s0022-5371(67)80149-x

Restom, K., Bangen, K. J., Bondi, M. W., Perthen, J. E., & Liu, T. T. (2007). Cerebral blood flow and BOLD responses to a memory encoding task: a comparison between healthy young and elderly adults. *Neuroimage, 37*(2), 430-439.

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, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926-1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (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, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The Necessity of the Medial Temporal Lobe for Statistical Learning. *Journal of Cognitive Neuroscience*, 1-12. doi: 10.1162/jocn_a_00578

Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (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

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003a). An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron, 37*(6), 1013-1025. doi: S0896627303001235 [pii]

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003b). Sequence? What Sequence?: the human medial temporal lobe and sequence learning. *Mol Psychiatry, 8*(11), 896-897. doi: 10.1038/sj.mp.4001424

Schwarb, H., & Schumacher, E. H. (2012). Generalized lessons about sequence learning from the study of the serial reaction time task. *Advances in cognitive psychology / University of Finance and Management in Warsaw, 8*(2), 165-178. doi: 10.2478/v10053-008-0113-1

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron, 60*(2), 378-389. doi: 10.1016/j.neuron.2008.09.023

Simon, S. S., Yokomizo, J. E., & Bottino, C. M. (2012). Cognitive intervention in amnestic Mild Cognitive Impairment: a systematic review. *Neurosci Biobehav Rev, 36*(4), 1163-1178. doi: 10.1016/j.neubiorev.2012.01.007

Szpunar, K. K., Addis, D. R., McLelland, V. C., & Schacter, D. L. (2013). Memories of the future: new insights into the adaptive value of episodic memory. *Frontiers in Behavioral Neuroscience, 7*, 47. doi: 10.3389/fnbeh.2013.00047

Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology-General, 134*(4), 552-564. doi: 10.1037/0096-3445.134.4.552

Tzvi, E., Muente, T. F., & Kraemer, U. M. (2014). Delineating the cortico-striatal-cerebellar network in implicit motor sequence learning. *Neuroimage, 94*, 222-230. doi: 10.1016/j.neuroimage.2014.03.004

Wang, Z., Jia, X., Liang, P., Qi, Z., Yang, Y., Zhou, W., & Li, K. (2012). Changes in thalamus connectivity in mild cognitive impairment: Evidence from resting state fMRI. *European Journal of Radiology, 81*(2), 277-285. doi: http://dx.doi.org/10.1016/j.ejrad.2010.12.044

Zhou, B., Liu, Y., Zhang, Z. Q., An, N. Y., Yao, H. X., Wang, P., . . . Jiang, T. Z. (2013). Impaired Functional Connectivity of the Thalamus in Alzheimer's Disease and Mild Cognitive Impairment: A Resting-State fMRI Study. *Current Alzheimer Research, 10*(7), 754-766.

Figures1-5

Click here to download Figure: luft_figures_rev.pdf

Figure 1

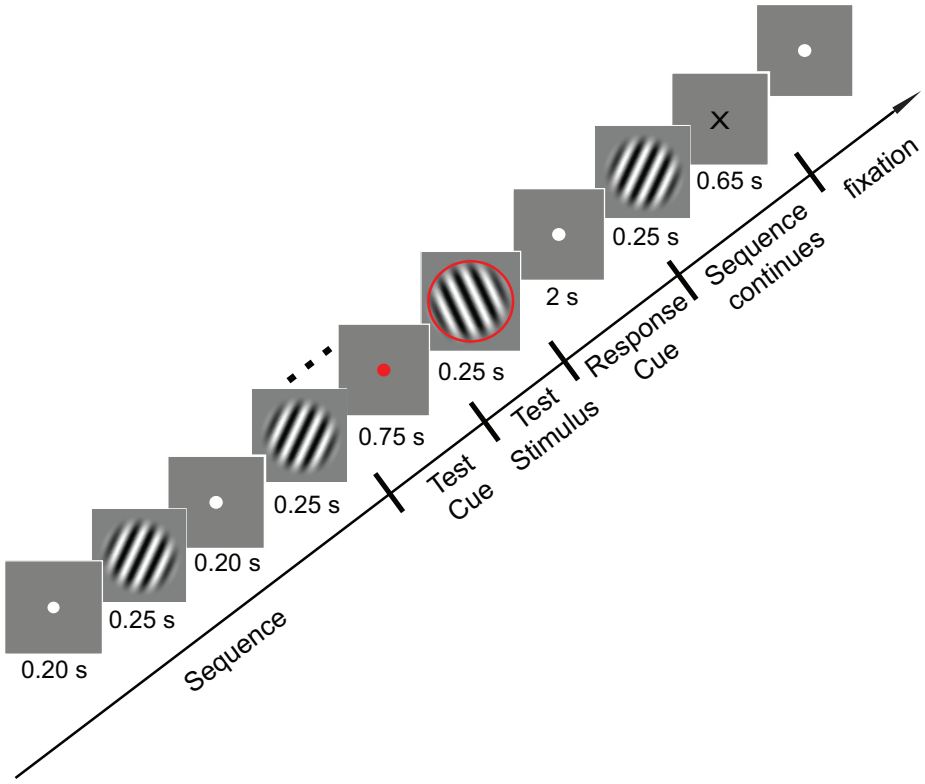
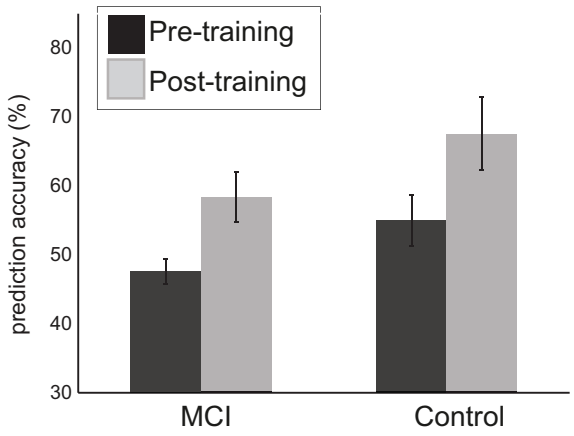


Figure 2

A. Behavioral Performance



B. Improvement Index

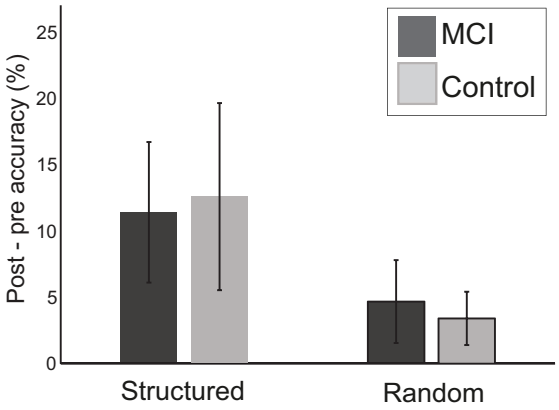


Figure 3

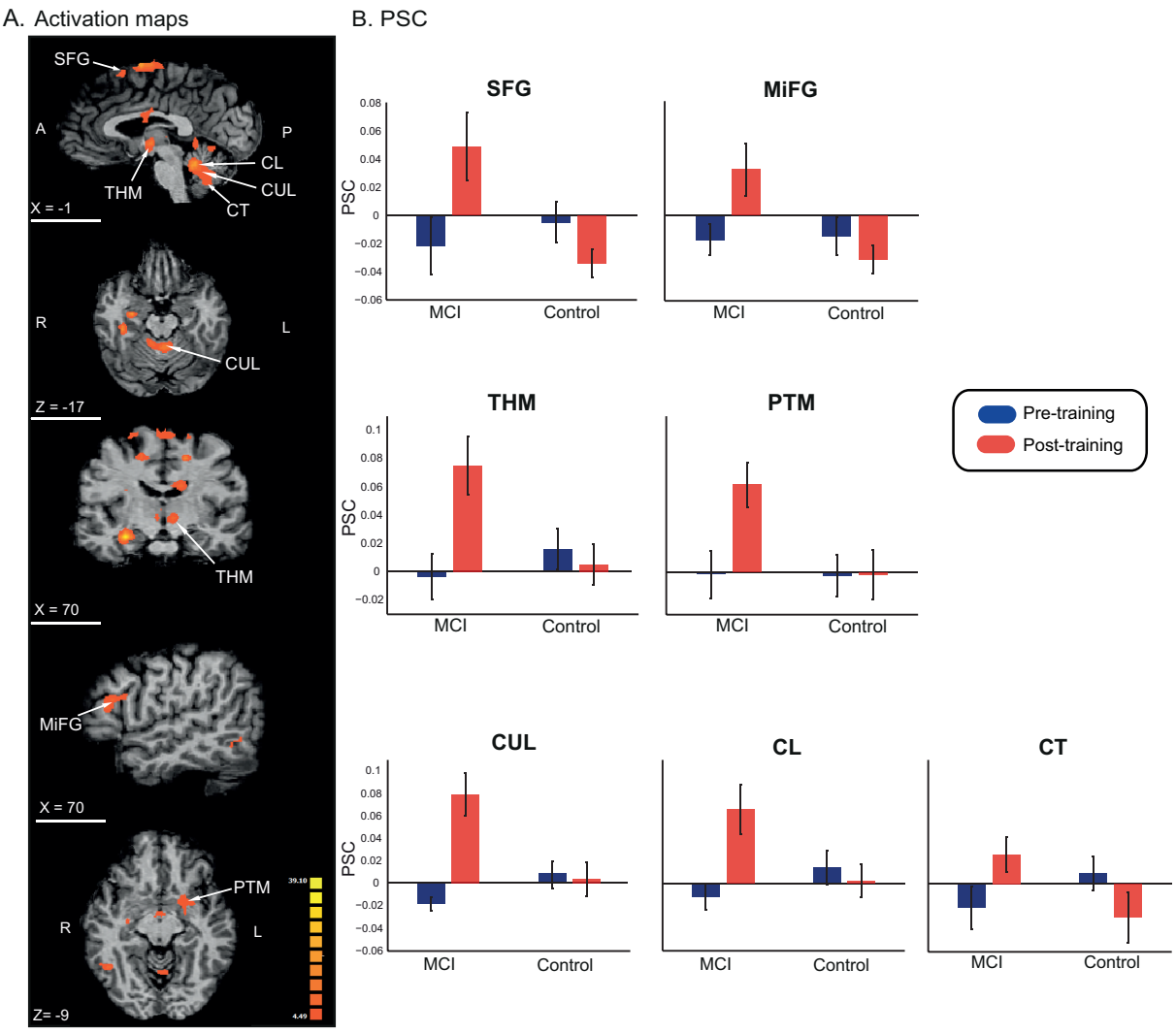

A. Activation maps
B. PSC
SFG
A
P
THM
CL
CUL
CT
X = -1
R
L
CUL
Z = -17
THM
X = 70
MiFG
X = 70
PTM
R
L
Z= -9
SFG
MiFG
THM
PTM
CUL
CL
CT
PSC
MCI
Control
Pre-training
Post-training

Figure 4

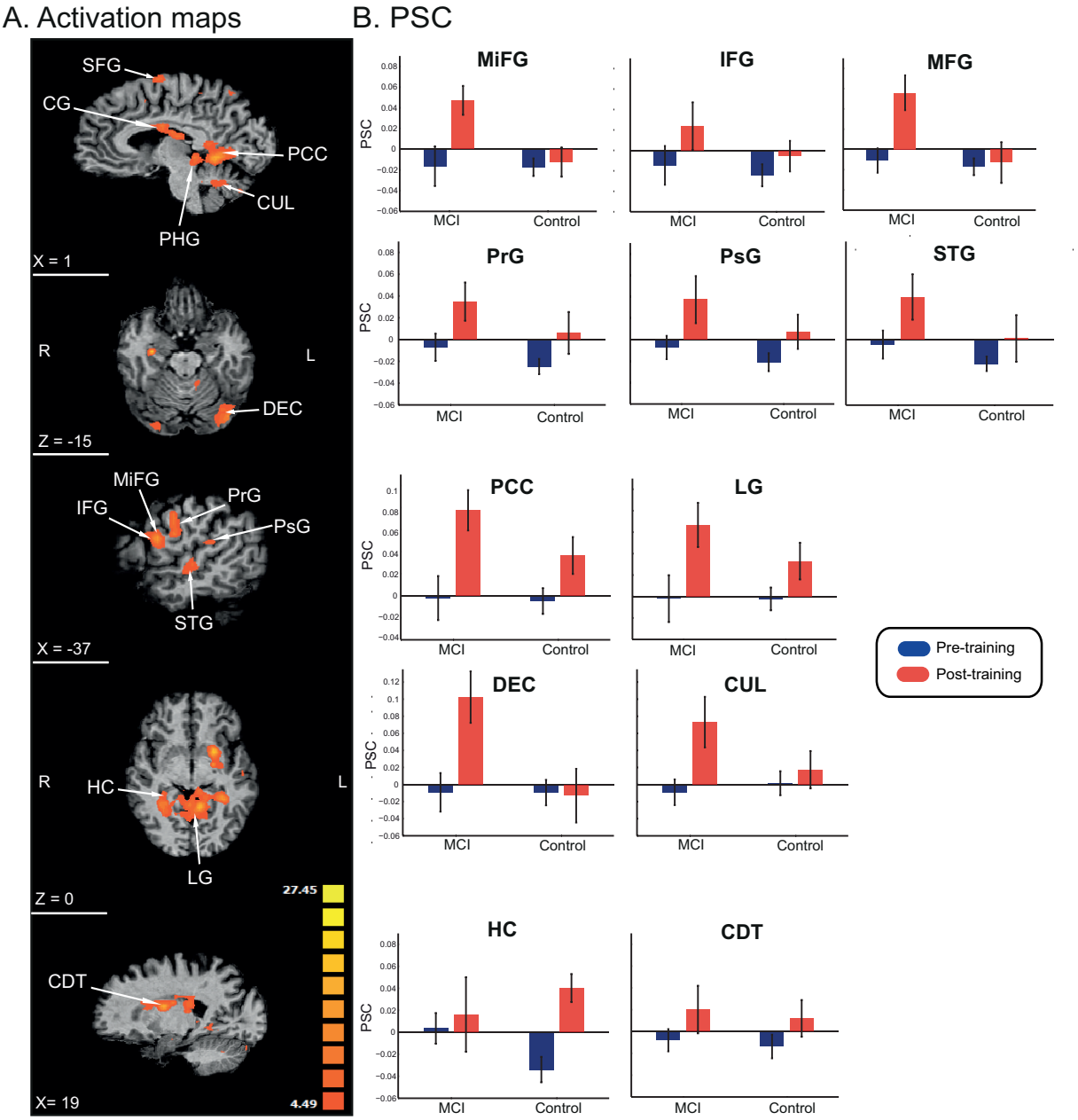
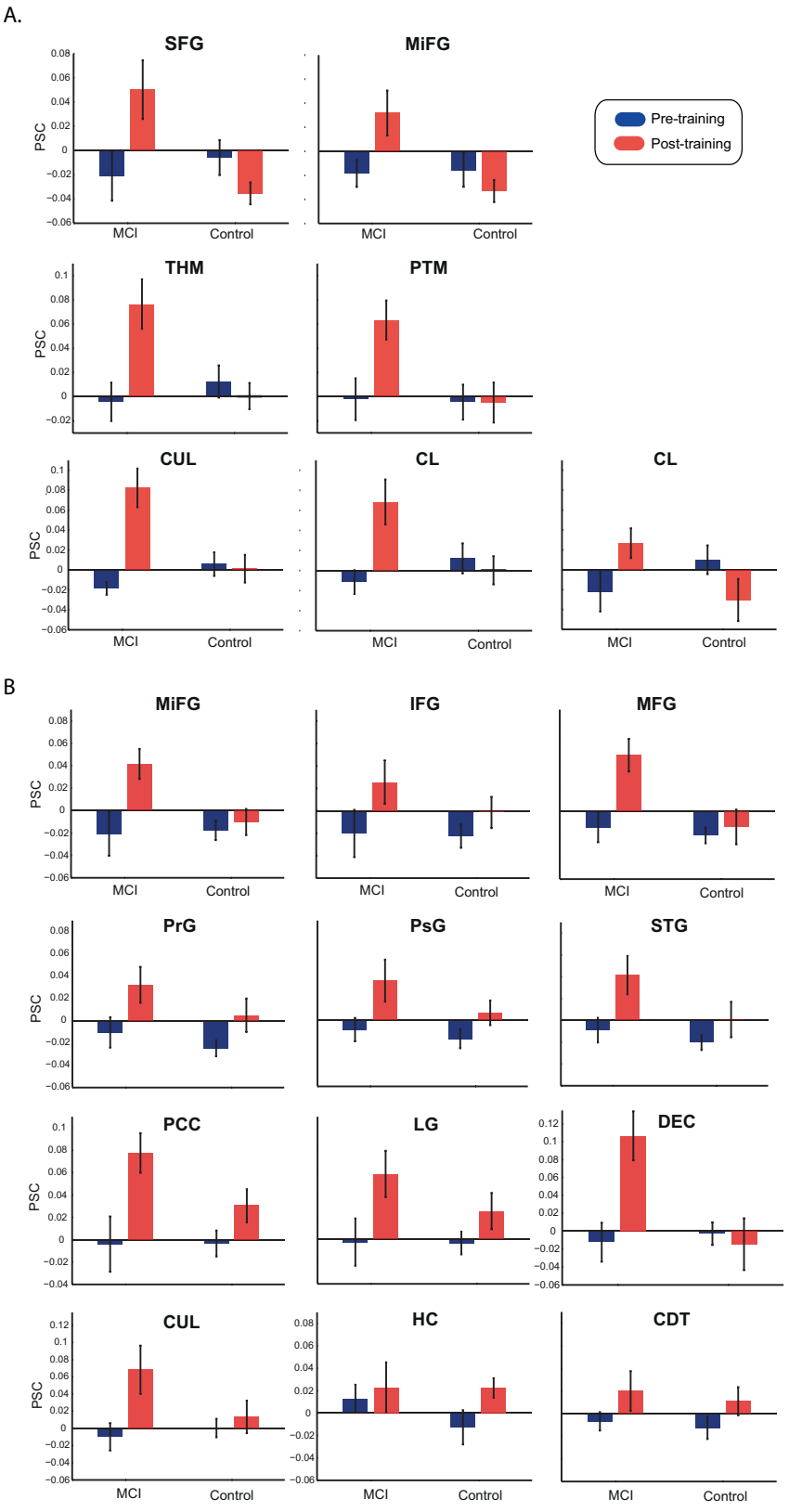

A. Activation maps
B. PSC
SFG
CG
PCC
CUL
PHG
X = 1
R
L
DEC
Z = -15
MiFG
IFG
PrG
PsG
STG
X = -37
R
HC
L
LG
Z = 0
27.45
CDT
X= 19
4.49
MiFG
IFG
MFG
PrG
PsG
STG
PCC
LG
DEC
CUL
HC
CDT
PSC
MCI
Control
Pre-training
Post-training

Figure 5


A.
SFG
MiFG
THM
PTM
CUL
CL
CL
B
MiFG
IFG
MFG
PrG
PsG
STG
PCC
LG
DEC
CUL
HC
CDT
PSC
MCI
Control
Pre-training
Post-training

Supplementary Material

Click here to download Supplementary Material: Luft_SI_rev2.pdf