Learning Temporal Statistics for Sensory Predictions in Aging

Caroline Di Bernardi Luft[1], Rosalind Baker[2], Aimee Goldstone[2], Yang Zhang[2], and Zoe Kourtzi[3]

Abstract

■ Predicting future events based on previous knowledge about the environment is critical for successful everyday interactions. Here, we ask which brain regions support our ability to predict the future based on implicit knowledge about the past in young and older age. Combining behavioral and fMRI measurements, we test whether training on structured temporal sequences improves the ability to predict upcoming sensory events; we then compare brain regions involved in learning predictive structures between young and older adults. Our behavioral results demonstrate that exposure to temporal sequences without feedback facilitates the ability of young and older adults to predict the orientation of an upcoming stimulus. Our fMRI results provide evidence for the involvement of corticostriatal regions in learning predictive structures in both young and older learners. In particular, we showed learning-dependent fMRI responses for structured sequences in frontoparietal regions and the striatum (putamen) for young adults. However, for older adults, learning-dependent activations were observed mainly in subcortical (putamen, thalamus) regions but were weaker in frontoparietal regions. Significant correlations of learning-dependent behavioral and fMRI changes in these regions suggest a strong link between brain activations and behavioral improvement rather than general overactivation. Thus, our findings suggest that predicting future events based on knowledge of temporal statistics engages brain regions involved in implicit learning in both young and older adults. ■

INTRODUCTION

Predicting upcoming events is critical for successful everyday interactions (e.g., avoiding obstacles, navigating in a new environment) and acquiring new skills (e.g., language and music). It is thought that the brain achieves this challenge by taking into account contextual associations that occur implicitly through exposure to the environment statistics. In particular, mere exposure to stimuli that co-occur in the environment is known to facilitate our ability to learn contingencies (for reviews, see Aslin & Newport, 2012; Perruchet & Pacton, 2006). For example, 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 (Turk-Browne, Junge, & Scholl, 2005; Fiser & Aslin, 2002; Chun, 2000; Saffran, Johnson, Aslin, & Newport, 1999; Saffran, Aslin, & Newport, 1996). This work suggests that observers acquire implicit knowledge of regularities and contextual associations, despite the fact that they may not be explicitly aware of its specific structure.

Previous neuroimaging studies have implicated subcortical and medial-temporal lobe regions in the learning of temporal statistics. In particular, the striatum and hippocampus have been implicated in learning of probabilistic associations (Shohamy & Wagner, 2008; Poldrack et al., 2001) and temporal sequences (Hsieh, Gruber, Jenkins, & Ranganath, 2014; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014; Schapiro, Kustner, & Turk-Browne, 2012; Gheysen, Van Opstal, Roggeman, Van Waelvelde, & Fias, 2011; Rose, Haider, Salari, & Buchel, 2011; Schendan, Searl, Melrose, & Stern, 2003; Rauch et al., 1997). For example, studies using the weather prediction paradigm (Foerde, Knowlton, & Poldrack, 2006; Poldrack et al., 2001; Poldrack, Prabhakaran, Seger, & Gabrieli, 1999) or the serial RT task (Aizenstein et al., 2006; Aizenstein, 2004; Schendan et al., 2003; Rauch et al., 1997) have implicated these regions in implicit learning of probabilistic associations. Although the hippocampus is thought to be involved at the initial stages of training, striatum and prefrontal regions have been shown to engage at later learning stages (Leaver, Van Lare, Zielinski, Halpern, & Rauschecker, 2009; Pasupathy & Miller, 2005; Poldrack & Packard, 2003). Furthermore, previous work has implicated mainly striatal regions in implicit learning (Hazeltine, Grafton, & Ivry, 1997; Rauch et al., 1997), whereas the medial-temporal lobe in both implicit and explicit learning (Dennis & Cabeza, 2011; Schendan et al., 2003).

Further work suggests that the same brain regions are involved in remembering the past and anticipating the future comprising medial frontoparietal and medial-temporal

[1]University of London, [2]University of Birmingham, [3]University of Cambridge

regions (Schacter, Addis, & Buckner, 2007). In particular, medial-temporal and parietal regions are thought to represent the spatiotemporal context in which events occur, whereas medial frontal regions use these contextual associations to make predictions about future events (Bar, 2009). This previous work has focused on explicit associative learning and episodic memories. However, the brain circuits that translate implicit knowledge of temporal statistics to future predictions remain largely unknown. Here, we test whether making predictions about the identity of upcoming stimuli based on implicit knowledge of temporal statistics engages frontostriatal circuits involved in implicit sequence learning rather than the medial frontotemporal circuits known to be primarily involved in explicit memory.

Furthermore, we test which brain regions support our ability to translate implicit knowledge to explicit predictions in older age. Previous behavioral studies on implicit learning of temporal regularities using sequence learning paradigms provide contradictory evidence: Some studies show little age-related impairment (Dennis, Howard, & Howard, 2006; Howard & Howard, 1992), whereas others show impaired performance in older adults when learning complex probabilistic sequences (Howard & Howard, 2013; Janacsek, Fiser, & Nemeth, 2012; Bennett, Howard, & Howard, 2007; Howard, Howard, Dennis, Yankovich, & Vaidya, 2004; Howard, Howard, Japikse, et al., 2004). In particular, young and older adults are shown to learn equally well early in the training, but older adults reach asymptotic performance at a lower level (Simon, Howard, & Howard, 2011; Howard, Howard, Dennis, & Kelly, 2008). Further brain imaging studies (Simon, Vaidya, Howard, & Howard, 2012; Dennis & Cabeza, 2011; Aizenstein et al., 2006; Raz et al., 2003; Ohnishi, Matsuda, Tabira, Asada, & Uno, 2001; Rubin, Van Petten, Glisky, & Newberg, 1999; Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998) show reduced activations in the striatum possibly because of structural decline as indicated by age-related changes in cortical thickness and white matter connectivity. These studies suggest that impairments in implicit learning of temporal sequences in aging may relate to striatal dysfunction (Howard & Howard, 2013). However, this previous work does not test how previous knowledge of temporal sequences may facilitate the explicit recognition of upcoming stimuli. Here, we combine behavioral and fMRI measurements to compare brain regions that are involved in translating implicit knowledge about temporal structures to explicit predictions in young and older adults.

In particular, we presented young and older participants with a sequence of leftward- and rightward-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. We show that the ability to predict the orientation of the test stimulus following exposure to structured sequences improved in
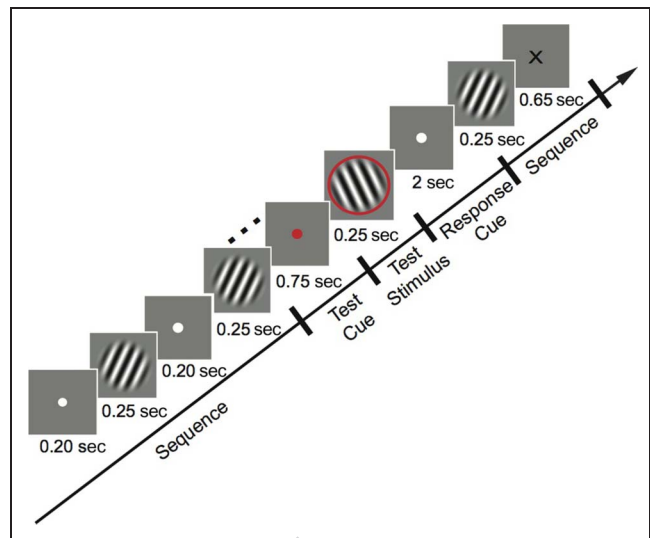

Figure 1. Design and behavioral performance. The trial design: A sequence of eight gratings was presented twice. A stimulus cue followed by a test grating was presented at a random position during the second repeat of the sequence. Following the response to the test stimulus, the remaining gratings of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

most young and older participants. Furthermore, our fMRI results provide evidence for the involvement of corticostriatal regions in learning predictive structures in both young and older learners. In particular, we observed learning-dependent fMRI responses for structured sequences in putamen for both young and older participants. However, young adults also showed learning-dependent fMRI responses for structured sequences in frontoparietal regions, whereas older adults mainly in thalamus. Our findings suggest that predicting future events based on the knowledge of temporal statistics engages brain regions involved in implicit learning in both young and older adults.

METHODS

Participants

Forty volunteers (20 young, 20 older adults) participated in this study. The young participants (11 women, 9 men) were recruited from the University of Birmingham (mean age = 24.1 years, SD = 2.0, range = 21–30). The older participants (11 men, 9 women) were recruited from the local community (mean age = 71.5 years, SD = 3.5, range = 65–78). One participant from the young group was excluded because of excessive head motion (>5 mm in most runs). All participants were naive to the aim of the study, had normal or corrected-to-normal vision, had no history of epilepsy or migraine, gave informed consent, and were paid for their participation. Older adults were screened using the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975): Scores were in the range of 27–30 indicating lack of cognitive impairment for this group. The study

was approved by the University of Birmingham Ethics Committee.

Stimuli

Stimuli comprised grayscale 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 local position adaptation because of stimulus repetition, we randomized the phase and jittered the grating orientation within a range of −2° to 2° across trials.

We used these stimuli to generate two sequences, each comprising eight gratings. Each grating orientation was presented four times in each sequence, and each sequence was repeated twice in a trial (the two repeats followed each other without gap), resulting in 16 stimuli presented successively per trial. The gratings in each trial were ordered, as shown below (1 refers to the leftward-oriented grating at −10° and 2 refers to the rightward-oriented grating at +10°):

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

Both Sequences A and B were intermixed (rather than blocked) and presented with equal probability across trials in a session. Intermixing the two sequences encouraged the participants to extract temporal regularities across trials rather than learn each sequence separately. Furthermore, 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 participants did not perform the task simply by memorizing the stimulus presented first in each trial, the orientation of the first stimulus was randomized in each trial; that is, for each of the two sequences, half of the trials started with leftward- and the rest with rightward-oriented gratings. Furthermore, 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, the two sequences were presented with the same probability (50%) across trials, 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 (i.e., temporal order associations among pairs or triplets of oriented gratings) rather than single temporal stimulus positions.

In addition to the 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. Random sequences were presented only before and after training (i.e., during scanning but not training sessions) to assess learning specificity to the trained structured sequences.

Stimuli were generated and presented using Psychtoolbox-3 (Brainard, 1997; Pelli, 1997). For the behavioral training sessions, stimuli were presented on a 21-in. CRT monitor (ViewSonic P225f 1280 × 1024 pixel, 85 Hz frame rate) at a distance of 45 cm. For the pretest and posttest fMRI scans, stimuli were presented using a projector and a mirror setup (1280 × 1024 pixel, 60 Hz frame rate) at viewing distance of 67.5 cm. 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

All the participants took part in two fMRI scans (pre- and posttraining) before and after behavioral training in the lab. Participants were familiarized with the task before scanning using simple repetitive sequences but were not trained on the sequences tested in the scanner. Most participants ($n = 8$) completed six training sessions ($n = 7$ completed five sessions; $n = 7$ completed four sessions; $n = 7$ completed three sessions) on consecutive days. The posttraining scan took place in the following day after the last behavioral training session.

Behavioral Training

For each trial, participants viewed 16 gratings (each sequence of eight gratings was repeated twice in a trial) presented sequentially on a gray background at the center of the screen. All stimuli were presented at the same rate; that is, each grating was presented for 0.3 sec followed by a fixation interval of 0.3 sec. Participants were asked to respond to a test stimulus that appeared for 0.3 msec surrounded by a red circle (0.3 sec). The test stimulus was preceded by a cue (red dot presented for 1 sec) and was followed by a white fixation (1700 msec). Participants were instructed to respond (the maximum RT allowed was 2000 msec), 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 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 sec), 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 runs of 40 trials each (20 per sequence type) with a minimum 2-min break between runs.

fMRI Design

Participants completed 6–10 runs of the prediction task without feedback in each scanning session (Figure 1). We aimed for eight runs per participant per session, but this was not always feasible within the time limits of the scanning sessions when participants experienced tiredness and discomfort or when they moved. Most participants (pre = 21, post = 18) completed eight runs; other participants had seven (pre = 7; post = 4) and six (pre = 1; post = 4) runs. Two participants had nine runs, and one participant had 10 runs in the posttraining scan session only. Comparing number of runs before versus after training (nonparametric Wilcoxon signed ranks test) did not show any significant differences across all participants ($Z = -.465$, $p = .642$), young ($Z = -.552$, $p = .581$) or older ($Z = 0$, $p = 1$) participants separately.

Each run lasted for 6 min and comprised five blocks of structured and five blocks of random sequences (three trials per block) presented in a random counterbalanced order. Each trial lasted for 10 sec, resulting in 30-sec long blocks. After every two stimulus blocks, a white fixation was presented for 10 sec. Each run started and ended with a fixation block. The trial design was similar to that used for behavioral training, that is, a structured sequence of eight gratings was presented twice in each trial. For random sequence trials, 16 left- or right-oriented gratings were presented in a random order. However, the presentation rate was faster during scanning than during behavioral training because of time constraints. That is, each grating was presented for 0.25 sec followed by fixation for 0.2 sec. The test stimulus appeared for 0.25 msec surrounded by a red circle (0.25 sec) and preceded by a cue (red dot, 0.75 sec). After the test stimulus, a fixation dot was presented for 2 sec (0.2 as red, remaining 1.8 sec as white), instructing the participants to respond whether the test matched (Button 1) the predicted stimulus or not (Button 2). After the response, the remaining gratings were presented till the end of the sequence, followed by an "X" cue (0.65 sec) marking the end of the trial.

fMRI Data Acquisition

fMRI data were acquired in a 3-T Achieva Philips scanner at the Birmingham University Imaging Centre using an eight-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; repetition time = 2 sec, echo time = 35 msec, flip angle = 73°, resolution = 2.5 * 2.5 * 4 mm,

no gap between slices, and ascending interleaved slice scan order).

Data Analysis

Behavioral Data Analysis

We assessed behavioral performance by accuracy (proportion 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

Preprocessing. Neuroimaging data were analyzed using Brain Voyager QX (Brain Innovation, Maastricht, the Netherlands). Preprocessing of functional data included slice scan time correction, three-dimensional motion correction, linear trend removal, temporal high-pass filtering (three cycles), and spatial smoothing using a 3-D Gaussian kernel of 5 mm FWHM. Blocks with head motion larger than 1 mm of translation or 1° of rotation 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.

General linear model. The BOLD responses to structured and random before and after training sequences were modeled 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 posttraining) 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 with a 30-sec (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. posttraining) × 2 (sequence: structured vs. random) × 2 (group: young vs. older). Statistical maps were cluster threshold corrected ($p < .005$, small cluster correction for 10 contiguous voxels; Forman et al., 1995).

Percent signal change analysis analysis. We identified brain regions across participants that showed significant learning-dependent changes in fMRI signals based on the whole-brain voxel-wise GLM analysis (RFX group analysis) described above. We then used these regions as anatomical ROIs of interest for further analyses of the fMRI responses (percent signal change [PSC]) to understand

differences between conditions (structured vs. random sequences) and groups (young vs. older) before and after training. We conducted this more elaborate analysis using a conservative split-half data procedure. 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 anatomically defined 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. This iterative procedure of selecting voxels within an ROI and evaluating their signals provides a conservative way of cross-validating our results.

In particular, for each participant, we defined ROIs based on the following interactions ($p < .01$): Session × Sequence and Session × Sequence × Group. To avoid circularity when selecting data from specific ROIs for further analysis, we defined ROIs based on half of the data (odd or even runs) and calculated PSC using the rest of the data. We split the data in odd and even runs and performed this procedure twice using either the odd or even data to define ROIs and compute the average PSC value from these two independent data samples. We considered ROIs that had volume higher than 300 mm^3 and showed a significant ($p < .05$) Session × Sequence or Session × Sequence × Age interaction following the split-half data procedure. Note that this lower threshold was adopted as the analysis was conducted using only half of the runs. 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.

Normalizing fMRI responses by vascular reactivity. To control for possible vascular differences between young and older participants, we normalized the PSC in each ROI by a vascular reactivity measure calculated based on fMRI responses to a w-hold task. Both young ($n = 6$) and older ($n = 13$) participants 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 sec) or breath normally when the white circle was presented (10 sec each). We collected data from one run lasting 200 sec (100 volumes including nine blocks for each condition plus one initial and one final 10-sec fixation period) for this task per participant. 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. Analysis of the fMRI data during this task showed a trend for higher hypercapnic responses in young than older adults, but no significant differences

between groups across ROIs (ROIs from Figure 3: $F(1, 16) = 0.07$, $p = .795$; ROIs from Figure 4: $F(1, 16) = 0.00$, $p = .994$).

RESULTS

Behavioral Results

We first tested participant improvement across training sessions. Only the first three sessions were considered
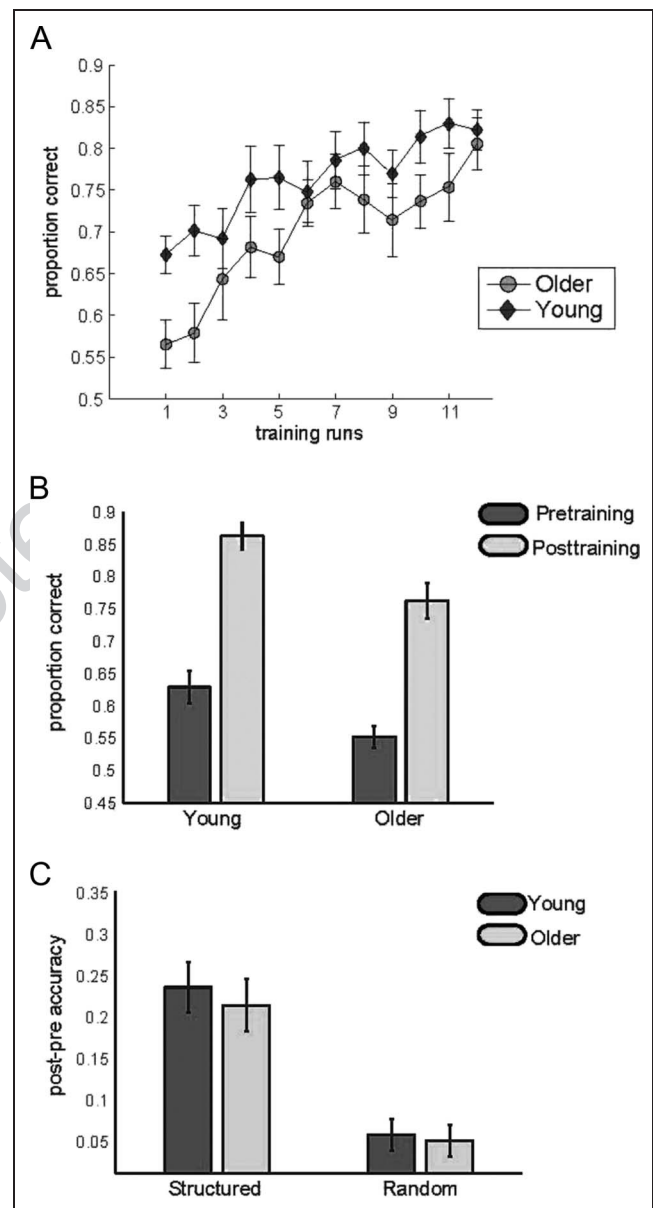

Figure 2. Behavioral performance. (A) Average performance (proportion correct) across participants for the structured sequences shown per training run for young and older participants. (B) Average performance (proportion correct) across participants for the structured sequences during the pre- and posttraining scanning sessions for young and older participants. (C) Performance improvement between scanning sessions was calculated for structured and random sequences by subtracting the average pretraining from the average posttraining performance during scanning.

in this analysis, as all participants completed training for at least three sessions. Including all participants in a 2 (group: young vs. older) × 3 (training sessions: Training Sessions 1, 2, 3) mixed design ANOVA showed a significant effect of training session ($F(1, 37) = 44.45, p < .001$), but no significant interaction of session and group ($F(1, 37) = .24, p = .621$), suggesting that performance improved across training for both young and older participants. We then repeated this analysis excluding participants who did not improve during training to allow direct comparison of brain activations between age groups that were not confounded by age-related performance differences (Hedden & Gabrieli, 2004). Most participants (16/20 young; 13/20 older) showed improved performance on the prediction task during training, as they gained more exposure to the temporal sequences (Figure 2A). Participants were excluded if they showed performance lower than 65% correct in the last training session and their performance did not change significantly across training runs compared with pretraining mean performance (that is, performance was lower than 1 standard deviation from the mean pretraining performance). A 2 (group: young vs. older) × 3 (training sessions: Training Sessions 1, 2, 3) mixed design ANOVA showed a sig-

nificant effect of session ($F(2, 54) = 39.60, p < .001$), but no significant interaction between session and group ($F(2, 54) = 1.03, p = .362$). Within-subject contrasts revealed a significant linear effect of session ($F(1, 27) = 53.09, p < .001$), but no significant interaction with group ($F(1, 27) = .41, p = .528$), suggesting that performance improved across training for both young and older participants. Furthermore, we observed a trend of better performance for younger than older participants, but this was not significant (main effect of group: $F(1, 27) = 2.94, p = .098$).

To compare performance for young and older learners before and after training, we conducted a three-way mixed design ANOVA (session: pre- vs. posttest; sequence: structured vs. random; group: young vs. older). Both young and older learners (Figure 2A) showed significant improvement in the task after training as indicated by a significant effect of session ($F(1, 27) = 111.8, p < .001$) and a nonsignificant interaction between session, sequence, and group ($F(1, 27) = 0.07, p = .791$). A significant main effect of sequence ($F(1, 27) = 190.7, p < .001$) and a significant Session × Sequence interaction ($F(1, 27) = 44.41, p < .001$) indicated better performance and higher improvement after training for structured than random
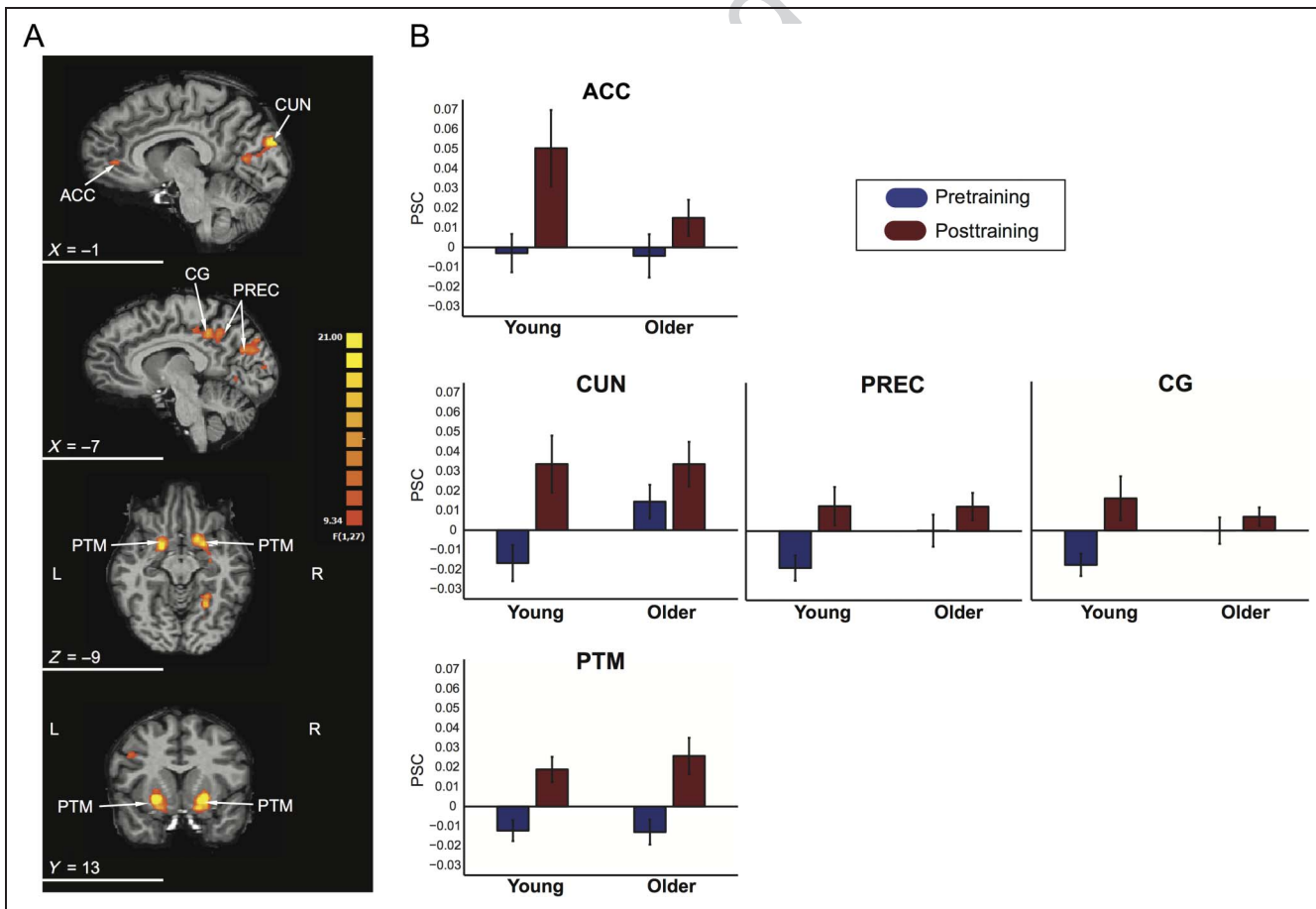

Figure 3. Learning-dependent fMRI changes: (A) GLM maps for the two-way interaction between session and sequence at $p < .005$ (small cluster threshold correction). (B) PSC for each of the ROIs identified by the two-way GLM interaction (Table 1) before and after training in young and older participants.

Figure 4. Comparing learning-dependent fMRI changes between age groups: (A) GLM maps for the three-way interaction between Session × Sequence × Group, at $p < .005$ (small cluster threshold correction). (B) PSC for each of the ROIs identified by the three-way GLM interaction (Table 2) before and after training in young and older participants. Performing the same analysis on all volunteers that participated in the study without excluding participants that showed weak improvement with training showed similar activation patterns.
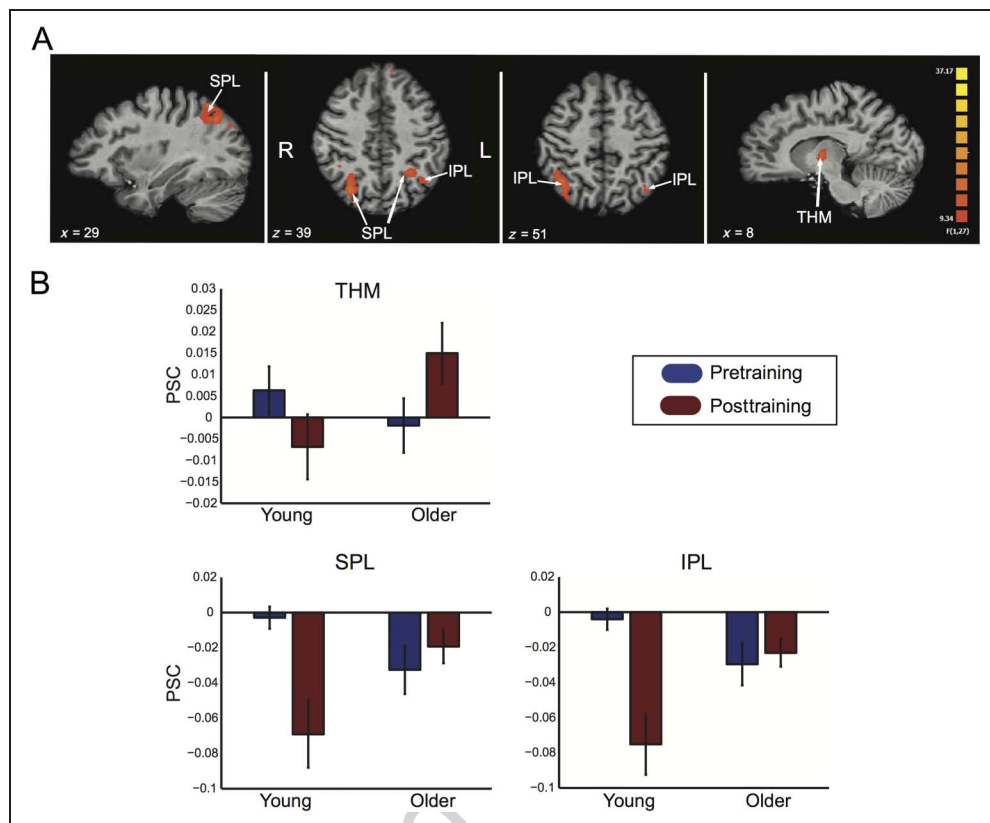
sequences. The lack of a significant interaction between session and group ($F(1, 27) = .298, p = .590$) suggests similar improvement in the prediction task for both young and older learners, despite or a nonsignificant but marginal main effect of group ($F(1, 27) = 3.59, p = .069$).

Furthermore, to quantify training-dependent improvement for the trained structured sequences, we computed an index subtracting pretraining from posttraining performance (Figure 2B). This improvement index did not differ significantly between young and older participants ($t(27) = .481, p = .634$). Thus, these analyses indicate similar training-dependent performance improvement for young and older learners that is specific to the trained structured rather than random sequences. That is, despite a nonsignificant trend for better performance in young adults, our statistical analyses show similar improvement across age groups allowing direct comparison of fMRI responses between young and older learners.

Finally, to control for differences in the number of training sessions across participants, we conducted three additional analyses. First, the number of training sessions was not significantly correlated with behavioral improvement ($r = -.274, p = .151$). Second, a mixed design ANOVA including the number of training sessions (3, 4, 5, 6) as a factor (session: pre- vs. posttest; sequence: structured vs. random; group: young vs. older; training session) showed similar results as the main analysis. That is, there was a significant effect of session ($F(1, 22) = 85.54, p < .001$), and sequence ($F(1, 22) = 132.0, p < .001$) and a significant Session × Sequence interaction ($F(1, 22) = 25.74, p < .001$) confirming higher improvement after training for structured than random sequences. Furthermore, the lack of a significant interaction between session and group ($F(1, 22) = 0.67, p = .421$) or a main effect of group ($F(1, 22) = 0.14, p = .0710$) suggests similar improvement in the prediction task for both young and older learners. Importantly, there was no effect of training sessions ($F(1, 22) = 1.20, p = .332$) nor any significant interaction with this factor ($p > .05$). Third, we compared behavioral performance during the pretraining session and the third training session, as all participants had completed this session. A 2 (session: pre- vs. last training) × 2 (group: young vs. older) ANOVA showed similar results as the main analysis; that is, a significant effect of session ($F(1, 27) = 19.94, p < .001$), but no significant effect of group ($F(1, 27) = 0.14, p = .0710$) nor a significant interaction between session and group ($F(1, 27) = 1.03, p = .320$). Taken together, these analyses suggest that it is unlikely that our results were significantly confounded by differences in training sessions across participants.

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 versus random sequences across sessions.

Table 1. Brain Regions Showing Significant Effects for the Session × Sequence Interaction ($p < .005$)

	H.	Volume (mm^3)	X	Y	Z	F	p
Anterior cingulate cortex	L	23	0	40	6	11.63	.002052
Anterior cingulate cortex	R	322	11	43	6	16.85	.000335
Precentral gyrus	R	1110	41	−14	36	36.06	.000002
Middle temporal gyrus	L	308	−46	−62	15	25.19	.000029
Cingulate gyrus	L	901	−7	−38	39	18.13	.000223
Precuneus	L	982	−1	−74	29	29.91	.000009
Precuneus	R	106	1	−74	27	21.93	.000071
Cuneus	L	2649	−1	−77	30	36.31	.000002
Cuneus	R	1835	2	−77	24	28.91	.000011
Parahippocampal gyrus	L	221	−25	−44	−9	22.28	.000065
Parahippocampal gyrus	R	128	17	0	−9	23.91	.000041
Putamen	L	3152	−19	10	0	36.43	.000002
Putamen	R	2006	20	13	0	39.75	.000001
Culmen	L	308	−22	−49	−9	21.22	.000088

A whole-brain GLM analysis (RFX, small cluster-threshold correction at $p < .005$) revealed a network of cortical and subcortical brain regions that showed differences in BOLD responses between structured and random sequences after than before training. We focus on the Session × Sequence interaction to test for brain regions showing training-induced improvement across participants (Table 1) and the three-way interaction of Group × Session × Sequence (Table 2) to identify regions showing differences between young and older participants in training-induced improvement. Furthermore, we conducted additional analyses to control for differences in the number of runs across participants. Most participants (pre = 21, post = 18) completed all eight runs per scanning session, and the number of runs was not significantly correlated with PSC before or after training across ROIs. Furthermore, we conducted the same GLM analysis including only

the first six runs per session, as all participants had at least six runs per session. We observed a similar pattern of activations as when including all runs, suggesting that it is unlikely that our results were confounded by differences in the number of runs across participants.

To interpret brain activations resulting from these interactions, for each participant and each of the regions identified from the whole-brain analysis, we extracted fMRI responses (PSC) for structured compared with random sequences before and after training. To avoid circularity in this ROI 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 ROIs 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

Table 2. Brain Regions Showing Significant Effects for the Session × Sequence × Group Interaction ($p < .005$)

	H.	Size	X	Y	Z	F	p
Superior frontal gyrus	L	335	−10	49	33	14.89	.000642
Thalamus	R	420	8	−5	6	19.13	.000164
Inferior parietal lobule	L	429	−40	−56	42	14.02	.000868
Inferior parietal lobule	R	1074	37	−52	51	31.33	.000006
Superior parietal lobule	L	66	−28	−48	39	17.13	.000306
Superior parietal lobule	R	916	35	−54	51	36.12	.000002
Precuneus	L	113	−28	−47	39	18.92	.000175
Precuneus	R	1026	26	−62	36	19.46	.000148

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 × Sequence or Session × Sequence × Group) for the split-half data procedure. As analysis of the behavioral data showed similar performance across runs within single scanning sessions, we averaged PSC signals across runs for each of the two (pre- and posttraining) sessions. In particular, comparing behavioral performance between the first (Runs 1–4) and second halves of scanning sessions did not show any significant differences. A 2 (session: pre vs. post) × 2 (time: first vs. second half) × 2 (group: young vs. older) mixed design ANOVA showed significant performance differences between pre- and posttraining sessions ($F(1, 27) = 107.11, p < .001$) but not between the first and second halves of the scanning sessions ($F(1, 27) = 1.57, p = .222$).

This PSC analysis revealed cortical and subcortical regions (Figure 3) that showed a significant session versus sequence interaction: medial frontal (anterior cingulate cortex), medial parietal (cuneus, precuneus, cingulate gyrus), and putamen. A two-session: (pre- vs. posttraining) × 2 (group: young vs. older) × ROI mixed design ANOVA showed increased fMRI responses for structured sequences after training (i.e., main effects of session) in medial frontal ($F(1, 27) = 5.80, p = .023$), medial parietal ($F(1, 27) = 14.07, p = .001$), and putamen ($F(1, 27) = 27.95, p < .001$). There was a significant interaction between Session × Group in medial parietal regions ($F(1, 27) = 9.24, p = .005$), but not in frontal regions ($F(1, 27) = 2.27, p = .143$) or putamen ($F(1, 27) = 0.181, p = .674$). In particular for young participants, we observed significantly increased fMRI responses for structured sequences after training (i.e., main effect of session) in medial frontal ($F(1, 15) = 5.56, p = .032$), medial parietal regions ($F(1, 15) = 19.13, p = .001$), and putamen ($F(1, 15) = 14.20, p = .002$). In contrast, for older participants, we observed increased fMRI responses for structured sequences after training in putamen ($F(1, 12) = 16.53, p = .002$), but not in the medial frontal ($F(1, 12) = 1.12, p = .310$) or medial parietal regions ($F(1, 12) = 0.40, p = .538$).

To further investigate differences in learning-dependent fMRI changes in young and older adults, we tested for regions that showed a significant interaction between Group × Session × Sequence. PSC analysis (Figure 4) showed significantly increased fMRI responses for structured sequences after training for older participants in thalamus. In contrast, for young participants, we observed a decrease in activation for structured sequences after training in lateral parietal areas (superior parietal lobe; inferior parietal lobe). These results were confirmed by a mixed ANOVA 2 (session: pre- vs. posttest) × 2 (group: young vs. older) that showed significant interactions between session and group in the thalamus ($F(1, 27) = 4.65, p = .040$) and parietal regions ($F(1, 27) = 12.61, p = .001$). These decreased activations after training possibly reflect adaptation to the trained structured sequences

or enhanced prediction error when sensory signals and top–down expectations do not correspond (Bastos et al., 2012; Summerfield & Egner, 2009; Friston, 2005) as in the case of unfamiliar random compared with the trained structured sequences. This activation decreases were evident in young but not older adults who did not show any significant learning-dependent decreased fMRI responses in parietal regions.

Furthermore, to test the link between learning-dependent behavioral and fMRI changes, we correlated behavioral improvement (performance post- minus pretraining) with fMRI activation differences between sessions (i.e., PSC post- minus pretraining). We tested for correlations of behavioral improvement across young and older adults with fMRI activation differences between sessions in regions that showed a significant Session × Sequence interaction (i.e., learning-dependent changes in fMRI responses for structured sequences independent of age). This analysis showed significant (Bonferroni corrected) correlations (Figure 5) in anterior cingulate ($r = .558, p < .001$), cingulate cortex ($r = .526, p = .008$), and the precuneus ($r = .492, p = .008$) across age groups. In contrast, correlations in putamen were significant for older ($r = .609, p = .032$) but not young ($r = .460, p = .376$) adults whereas correlation in inferior parietal lobe were significant for young ($r = -.683, p = .008$) but not older ($r = .201, p = 1$) adults. Note that correlations were positive in frontal and striatal regions consistent with enhanced activations after training in these regions whereas negative in parietal regions consistent with decreased activations in young adults after training. These correlations suggest a link between behavioral improvement and learning-dependent fMRI changes—rather than general overactivation—in regions shown to be involved in learning predictive temporal structures.

We then tested whether enhanced fMRI responses for structured sequences after training could be explained by differences in the fMRI response to random sequences. Analyzing PSC data from fixation baseline for random sequences did not show any significant interactions between session and group in medial frontal (anterior cingulate cortex: $F(1, 27) = 0.163, p = .690$), medial parietal ($F(1, 27) = 0.012, p = .914$), putamen ($F(1, 27) = 0.91, p = .347$), thalamus ($F(1, 27) = 0.665, p = .800$), or lateral parietal regions ($F(1, 27) = 0.56, p = .559$). Furthermore, there were no significant effects for session for medial frontal ($F(1, 27) = 3.44, p = .074$), thalamus ($F(1, 27) = 0.635, p = .433$), or lateral parietal ($F(1, 27) = 0.15, p = .705$) and only marginal effects in medial parietal ($F(1, 27) = 4.29, p = .048$) and putamen ($F(1, 27) = 4.44, p = .045$). 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 versus after training.

Finally, consistent with our previous studies (Baker, Dexter, Hardwicke, Goldstone, & Kourtzi, 2014), we observed variability in learning-dependent improvement in
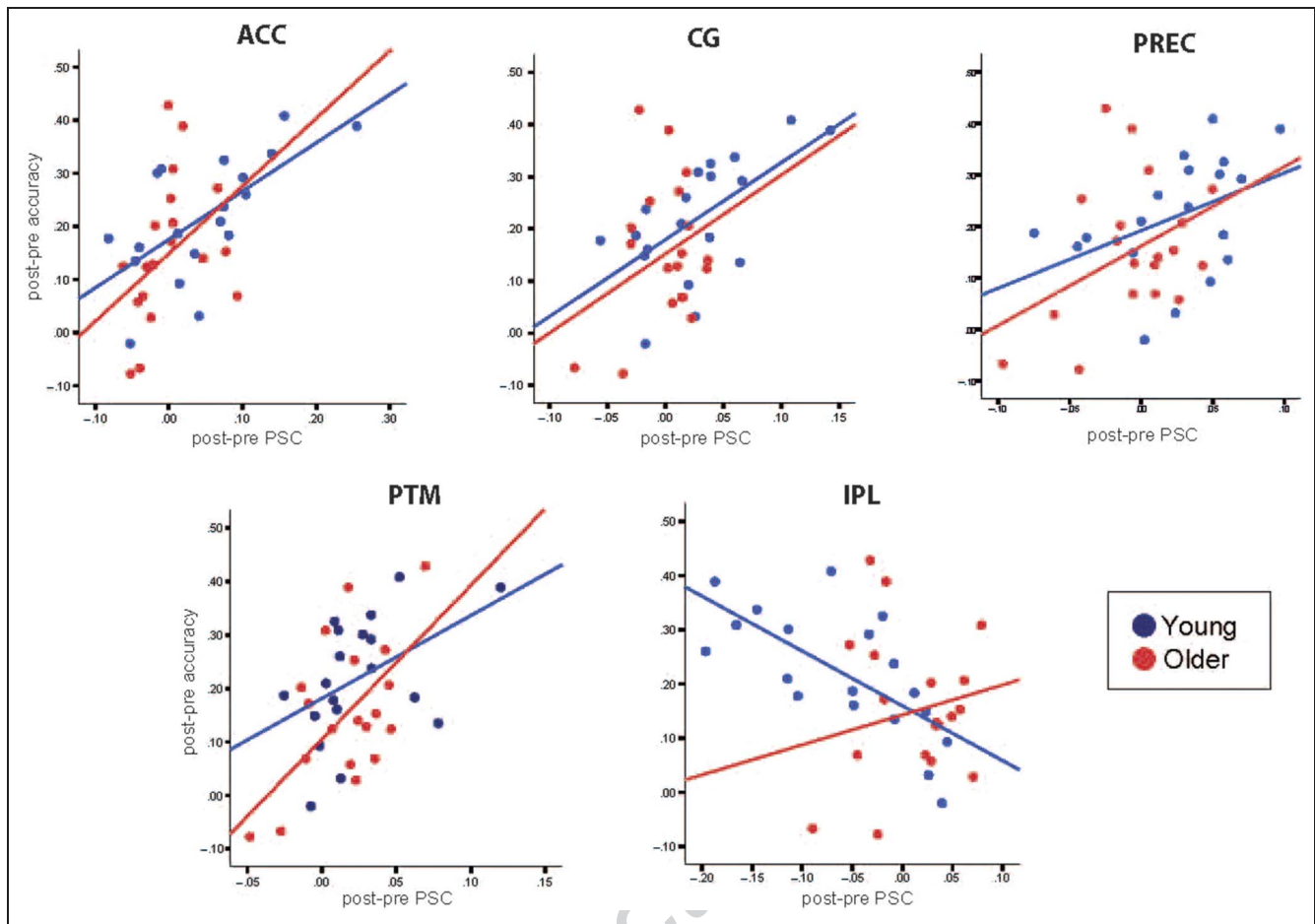
Figure 5. Correlations between behavioral and brain changes. Correlations (Bonferroni-corrected) between behavioral improvement (accuracy in the posttraining minus pretraining scanning session) and PSC (PSC in the posttraining minus pretraining scanning session). The following ROIs showed significant correlations: ACC = anterior cingulate cortex, CG = cingulate gyrus, PREC = precuneus, PTM = putamen, IPL = inferior parietal lobe. Regression lines are shown only for each group (blue for young, red for older adults).

both young and older adults. That is, although most participants improved across training sessions, some did not show enhanced performance during training. Previous studies have shown different activation patterns in high- versus low-performing older adults (Hedden & Gabrieli, 2004). To ensure that any differences in activation patterns between age groups were not confounded by differences in performance, we excluded weaker learners from both groups (i.e., participants whose performance remained below 65% correct and did not change significantly from the pretraining session across training sessions). Furthermore, GLM analysis including all participants (i.e., without excluding participants that did not improve during training) showed similar activation patterns, suggesting that our results were not biased by excluding weaker learners.

Control Analyses

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

First, we tested whether the participants used a simple heuristic strategy to solve the task. For example, it is possible that the participants detected that there were more alternation (12, 21) than repetition pairs (11, 22) in the structured sequences. We reasoned that, if the participants used this alternation heuristic, they would show higher proportion of responses for alternation pairs independent of the sequence (i.e., structured or random). In particular, we calculated the probability with which each possible pair (11, 12, 21, 22) appeared and the participant responses per pair comprising the test grating and its preceding item in the sequence. This analysis showed that participants responded similarly across alternation and repetition pairs in random sequences. In contrast, for structured sequences the participants' responses followed the frequency with which alternation versus repetition pairs were presented. That is, participants detected both the alternation and the less frequent repetition pairs. This result was supported by a 3 (sequence type: A, B, Random) $\times$ 2 (session: pre- vs. posttraining) $\times$ 4 (pair: 11,12,21,22) that showed a significant three-way interaction between Sequence Type $\times$ Session $\times$ Pair ($F(6, 168) = 4.50$, $p < .001$) and significant two-way interactions

between Sequence Type × Pair: $F(6, 168) = 12.99$, $p < .001$) and Pair × Session ($F(3, 84) = 3.83$, $p = .013$). Thus, this analysis suggests that participants extracted temporal statistics across items specific to the trained sequences rather than applying a single heuristic strategy (i.e., detect alternation and ignore repetition pairs) to solve the task.

Second, 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 sequences were structured and some random. Furthermore, comparing RTs to structured and random sequences in the pre- and posttraining session between the two groups (three-way mixed design ANOVA: Session × Sequence × Group) showed decreased RTs after training (main effect of session: $F(1, 27) = 23.79$, $p < .001$), but no significant differences between structured and random sequences (main effect of sequence ($F(1, 27) = 0.08$, $p = .783$), 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 ($F(1, 27) = 0.41$, $p = .841$) nor a three-way interaction (Session × Sequence × Group, $F(1, 27) = 0.12$, $p = .735$), suggesting that our results could not be simply due to age-related differences in processing speed, sensorimotor performance, or attention (Ren, Wu, Chan, & Yan, 2013; Birren & Fisher, 1995).

Third, it is possible that the differences in activation patterns observed between young and older participants originated from differences in vascular reactivity rather than differences in underlying neuronal activity (Restom, Bangen, Bondi, Perthen, & Liu, 2007; D'Esposito, Deouell, & Gazzaley, 2003; Hamzei, Knab, Weiller, & Rother, 2003; D'Esposito, Zarahn, Aguirre, & Rypma, 1999). To control for this possibility, we acquired fMRI data in both young and older participants 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 10-sec breath-holding vs. normal breathing periods) to normalize the stimulus-evoked BOLD signal, as previously described (Handwerker et al., 2007). For each ROI, 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 6) showed a similar pattern of results as in Figure 3 (Figure 6A) and Figure 4 (Figure 6B), suggesting that the differences in activation patterns that we observed between young and older adults could not be simply explained by differences in vascular reactivity between participant groups. Furthermore, 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 in these regions demonstrates that we recorded fMRI signals with similar sensitivity in young and older participants, allowing us to compare between brain regions and participant groups. Signal-to-noise ratio was calculated as (Stask – Sfixation)/std(Stask, Sfixation), where Stask indicates BOLD signal during the task blocks including structured and random sequences and Sfixation indicates BOLD signal during the fixation periods. The difference of the signal between task and fixation was then divided by the standard deviation of the BOLD signal of the whole time series including both task and fixation. In particular, a 2 (session: pre- vs. post-training) × 2 (group: young vs. older) mixed design ANOVA did not show any significant interactions between session and group in medial frontal ($F(1, 27) = 2.59$, $p = .119$), medial parietal ($F(1, 27) = 0.06$, $p = .814$), putamen ($F(1, 27) = 0.19$, $p = .668$), thalamus ($F(1, 27) = 0.21$, $p = .650$), or lateral parietal ($F(1, 27) = 3.05$, $p = .092$).

Fourth, we tested whether differences in brain activations between age groups were due to baseline differences in performance before training. Our behavioral results showed similar behavioral improvement for young and older participants (as shown by lack of a significant interaction between group and session) allowing direct comparison of fMRI activity patterns between age groups. Furthermore, correlating pretraining behavioral performance with fMRI responses (PSC index: PSC posttraining minus PSC pretraining) did not show any significant correlations in regions ($p > .05$) showing training-dependent changes. Thus, these results suggest that differences in brain activation patterns between age groups could not be simply attributed to differences in task performance before training.

Finally, we consider possible limitations of the block fMRI design adopted in our study. In particular, we opted for a block rather than an event-related fMRI design, as the long trials (10 sec per trial)—necessitated by the sequence presentation—would result in the acquisition of small numbers of trials per session. In this case, an event-related design would be most likely underpowered. Therefore, we chose to block a small number of trials (three trials per block, 30 sec per block) to ensure that our experimental design would afford sufficient detection power. One possible limitation of block fMRI designs is that they may result in differences in the participants' attentional state across conditions or encourage the participants to adopt heuristic strategies for solving the task. Our control analyses—as described above—suggest that these alternate interpretations are rather unlikely. First, comparing RTs to structured and random sequences in the pre- and posttraining session between the two groups showed that participants engaged with the task when both structured and random sequences were presented. Second, any heuristics followed by the participants (e.g., alternation detection strategy) would be expected to be more prominent after than before training. As a result, if the
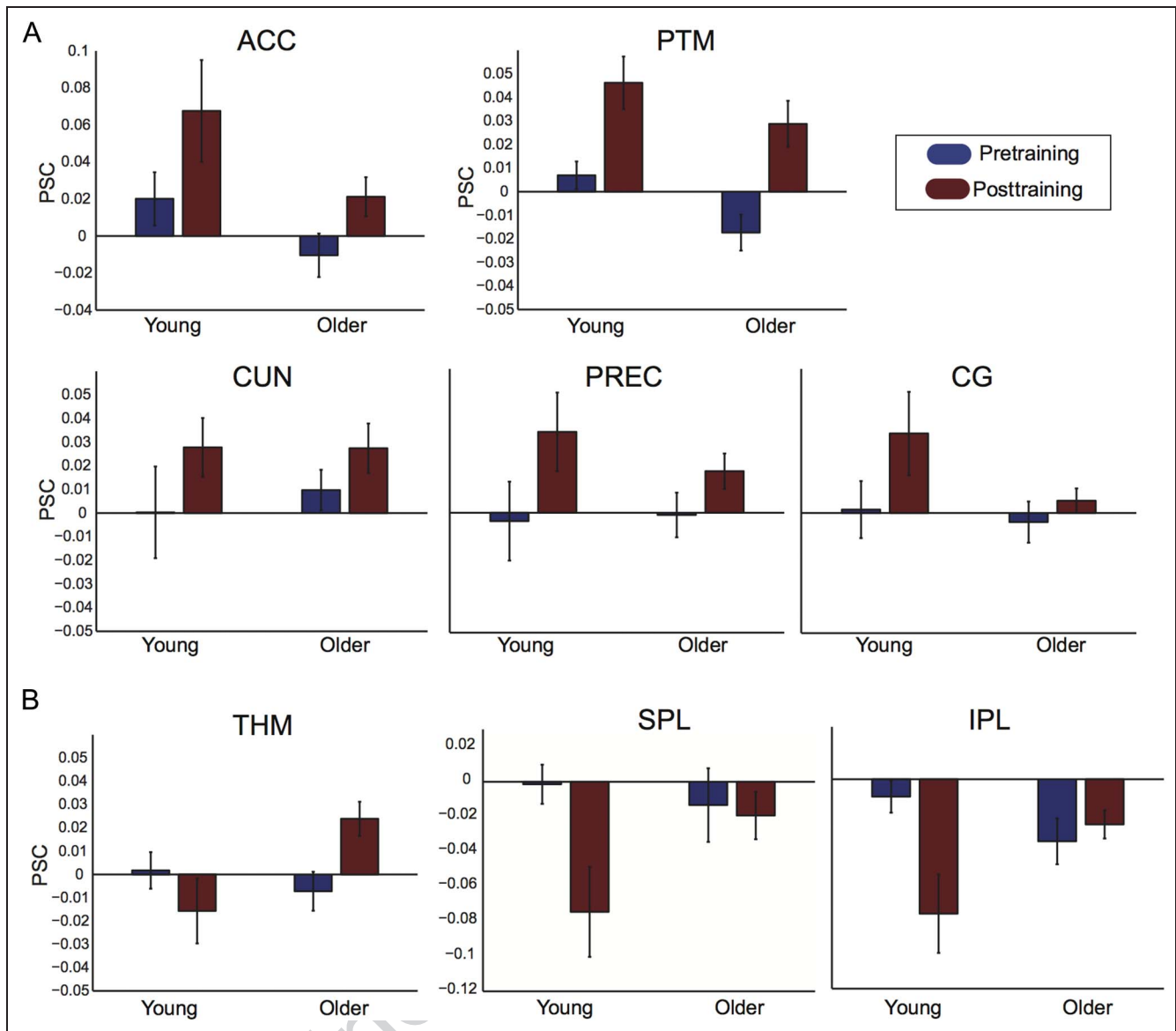
Figure 6. fMRI responses normalized by vascular reactivity. For each ROI, we used differences in the fMRI response between breath-holding and the normal breathing periods to normalize the PSC for structured and random sequences. Data are shown for ROIs that had more than 300 mm³ and showed a significant interaction between session (pre- vs. posttraining) sequence (structured vs. random) and group (young vs. older) for the split-half data procedure.

participants applied such strategies to the blocks of random sequences, we would expect differences in the fMRI responses to random stimuli before versus after training. In contrast to this prediction, we did not observe any significant interactions between session and group ($p < .05$) nor significant effects of session ($p < .05$) for fMRI responses to random stimuli. Furthermore, debriefing the participants after training indicated that the participants were not aware that some sequences were structured and some random, suggesting that it is unlikely that fMRI responses to structured sequences were due to the fact that participants ignored the random blocks. Finally, the block design does not allow us to separate correct from incorrect trials. However, both young and older participants performed on average higher than 75% correct after

training, resulting in a small number of incorrect trials. Thus, it is unlikely that differences in activation patterns between young and older adults could be driven by the small number of incorrect trials after training. Further significant correlations between brain activations and behavioral performance for both young and older adults suggest that the observed activation patterns relate to the participants' leaning performance.

DISCUSSION

Our results demonstrate that exposure to temporal sequences facilitates the ability to predict the identity of upcoming events in both young and older age. Most young and older participants improved after training in the

prediction task, allowing us to compare directly brain activity between the two groups. Our fMRI results provide evidence for the involvement of corticostriatal regions in learning predictive structures in both young and older learners. In particular, for young adults, we observed learning-dependent fMRI changes in frontoparietal and striatal regions. However, for older adults, learning-dependent activations were observed mainly in subcortical (putamen, thalamus) regions but were weaker in frontoparietal regions. These results are consistent with the role of frontoparietal regions (Aizenstein, 2004; Sakai et al., 1998) and the striatum (Berns, Cohen, & Mintun, 1997; Rauch et al., 1997) in implicit sequence learning. Thus, our findings suggest that predicting future events based on knowledge of temporal statistics engages regions involved in implicit learning in both young and older adults.

Our findings advance our understanding of the neural mechanisms that support learning to predict the identity of upcoming stimuli based on the knowledge of temporal structures in two main respects. First, our study tests the role of sequence learning on explicit predictive judgments related to visual recognition in aging. Previous work on learning temporal sequences has focused on implicit measures of sequence learning, such as familiarity judgments or RTs. For example, the Serial Reaction Time Task (Nissen & Bullemer, 1987; for a 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 predetermined sequence, and participants are asked to press the corresponding keys. Training results in faster RTs for trained than for random sequences, indicating anticipatory behavioral effects. Although these findings suggest that we are prepared to react faster to anticipated stimuli, these paradigms do not allow us to test how previous knowledge of temporal sequences may facilitate the explicit recognition of upcoming stimuli. Here, using an explicit prediction test, we demonstrate that predictions related to identification of objects are facilitated by knowledge of temporal structures.

Second, our fMRI results showing increased striatal activation for structured sequences in both young and older adults appear to contradict neuroimaging evidence for structural striatal decline (i.e., cortical thickness) with age (Fjell et al., 2013; Raz et al., 2003). In particular, previous studies suggest that behavioral impairment in sequence learning in aging may relate to this striatal dysfunction (Howard & Howard, 2013). However, despite this structural decline, functional imaging studies using the weather prediction paradigm or the serial RT task show compensatory brain activity in older adults (i.e., higher activation in older than young participants) that may mediate successful performance in these tasks (Dennis & Cabeza, 2008; Cabeza, Anderson, Locantore, & McIntosh, 2002). Such activity is observed in the hippocampus (Simon et al., 2012; Dennis & Cabeza, 2011), parietal cortex (Fera

et al., 2005), and striatum (Rieckmann & Backman, 2009; Aizenstein et al., 2006). Our results show that—despite a nonsignificant trend for better performance in young adults—both young and older adults showed similar learning-dependent improvement in the prediction task. Furthermore, older learners (i.e., participants that improved during training) showed learning-dependent activity (i.e., higher activation for structured than random sequences after training) in putamen and thalamus, suggesting that older learners may rely mainly on subcortical regions for translating implicit knowledge to predictions of upcoming events. Importantly, correlations of behavioral improvement with enhanced striatal activity in older adults suggest a link between brain activation and learning in the striatum rather than general overactivation related to neural inefficiency because of age-related decline.

It is possible that the enhanced involvement of subcortical regions in learning that we observed for older adults is due to training across multiple sessions. Previous work has implicated the hippocampus in rapid associative learning early in training while striatal regions in the integration of probabilistic structures over extended temporal intervals (Poldrack & Packard, 2003). Further work has shown that, although both young and older adults recruit hippocampal regions early in training, only young adults recruit striatal regions later in training, whereas older adults continue to recruit the hippocampus for sequence learning (Simon et al., 2012). These findings suggest dedifferentiation of hippocampal learning systems in older age for short-term training (i.e., single session). In contrast, our findings suggest that prolonged training over multiple sessions may result in striatal involvement in implicit learning in older adults. Thus, it is possible that the temporal window for integrating probabilistic structures in the striatum may extend across multiple training sessions.

Finally, similar to our previous studies (Baker et al., 2014), we observed individual variability in the ability to predict upcoming events based on implicit knowledge of temporal structures. That is, although most participants improve during training across sessions, some do not show enhanced performance with training. Understanding the sources of this individual variability requires larger-scale studies that investigate the neurocognitive abilities that may contribute to individual differences in learning ability across larger population samples. Furthermore, we have previously (Baker et al., 2014) shown that improvement in the prediction task lasted for a prolonged period (up to 3 months). Future work is needed to investigate whether longer-term training may result in stronger and longer-lasting brain plasticity for learning predictive structures across the lifespan.

REFERENCES

Aizenstein, H. J. (2004). Regional brain activation during concurrent implicit and explicit sequence learning. *Cerebral Cortex, 14,* 199–208.

Aizenstein, H. J., Butters, M. A., Clark, K. A., Figurski, J. L., Andrew Stenger, V., Nebes, R. D., et al. (2006). Prefrontal and striatal activation in elderly subjects during concurrent implicit and explicit sequence learning. *Neurobiology of Aging, 27,* 741–751.

Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science, 21,* 170–176.

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 Research, 99,* 124–133.

Bar, M. (2009). The proactive brain: Memory for predictions. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 364,* 1235–1243.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron, 76,* 695–711.

Bennett, I. J., Howard, J. H., Jr., & Howard, D. V. (2007). Age-related differences in implicit learning of subtle third-order sequential structure. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 62,* P98–P103.

Berns, G. S., Cohen, J. D., & Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science, 276,* 1272–1275.

Birren, J. E., & Fisher, L. M. (1995). Aging and speed of behavior: Possible consequences for psychological functioning. *Annual Review of Psychology, 46,* 329–353.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *Neuroimage, 17,* 1394–1402.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4,* 170–178.

Dennis, N. A., & Cabeza, R. (2008). Neuroimaging of healthy cognitive aging. *The handbook of aging and cognition.* New York: Psychology Press.

Dennis, N. A., & Cabeza, R. (2011). Age-related dedifferentiation of learning systems: An fMRI study of implicit and explicit learning. *Neurobiology of Aging, 32,* 2318.e17–2318.e30.

Dennis, N. A., Howard, J. H., & Howard, D. (2006). Implicit sequence learning without motor sequencing in young and old adults. *Experimental Brain Research, 175,* 153–164.

D'Esposito, M., Deouell, L. Y., & Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: A challenge for neuroimaging. *Nature Reviews Neuroscience, 4,* 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,* 6–14.

Fera, F., Weickert, T. W., Goldberg, T. E., Tessitore, A., Hariri, A., Das, S., et al. (2005). Neural mechanisms underlying probabilistic category learning in normal aging. *Journal of Neuroscience, 25,* 11340–11348.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology-Learning Memory and Cognition, 28,* 458–467.

Fjell, A. M., Westlye, L. T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., et al. (2013). Critical ages in the life course of the adult brain: Nonlinear subcortical aging. *Neurobiology of Aging, 34,* 2239–2247.

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 11778–11783.

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

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine, 33,* 636–647.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 360,* 815–836.

Gheysen, F., Van Opstal, F., Roggeman, C., Van Waelvelde, H., & Fias, W. (2011). The neural basis of implicit perceptual sequence learning. *Frontiers in Human Neuroscience, 5,* 137.

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,* 1393–1399.

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. *Human Brain Mapping, 28,* 846–859.

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

Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience, 5,* 87–96.

Howard, D. V., & Howard, J. H. (1992). Adult age differences in the rate of learning serial patterns—Evidence from direct and indirect tests. *Psychology and Aging, 7,* 232–241.

Howard, D. V., Howard, J. H., Jr., Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit sequence learning: Effects of level of structure, adult age, and extended practice. *Psychology and Aging, 19,* 79–92.

Howard, J. H., Howard, D. V., Dennis, N. A., & Kelly, A. J. (2008). Implicit learning of predictive relationships in three-element visual sequences by young and old adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1139–1157.

Howard, J. H., Howard, D. V., Dennis, N. A., Yankovich, H., & Vaidya, C. J. (2004). Implicit spatial contextual learning in healthy aging. *Neuropsychology, 18,* 124–134.

Howard, J. H., Jr., & Howard, D. V. (2013). Aging mind and brain: Is implicit learning spared in healthy aging? *Frontiers in Psychology, 4,* 817.

Hsieh, L. T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron, 81,* 1165–1178.

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science, 15,* 496–505.

Leaver, A. M., Van Lare, J., Zielinski, B., Halpern, A. R., & Rauschecker, J. P. (2009). Brain activation during anticipation of sound sequences. *Journal of Neuroscience, 29,* 2477–2485.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning—Evidence from performance measures. *Cognitive Psychology, 19,* 1–32.

Ohnishi, T., Matsuda, H., Tabira, T., Asada, T., & Uno, M. (2001). Changes in brain morphology in Alzheimer disease and normal aging: Is Alzheimer disease an exaggerated aging process? *American Journal of Neuroradiology, 22,* 1680–1685.

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233–238.

Poldrack, R. A., Clark, J., Pare-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., et al. (2001). Interactive memory systems in the human brain. *Nature, 414,* 546–550.

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia, 41,* 245–251.

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

Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., Brown, H. D., et al. (1997). Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Human Brain Mapping, 5,* 124–132.

Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging: Evidence from structural magnetic resonance imaging. *Neuropsychology, 12,* 95–114.

Raz, N., Rodrigue, K. M., Kennedy, K. M., Head, D., Gunning-Dixon, F., & Acker, J. D. (2003). Differential aging of the human striatum: Longitudinal evidence. *AJNR, American Journal of Neuroradiology, 24,* 1849–1856.

Ren, J., Wu, Y. D., Chan, J. S., & Yan, J. H. (2013). Cognitive aging affects motor performance and learning. *Geriatrics & Gerontology International, 13,* 19–27.

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,* 430–439.

Rieckmann, A., & Backman, L. (2009). Implicit learning in aging: Extant patterns and new directions. *Neuropsychology Review, 19,* 490–503.

Rose, M., Haider, H., Salari, N., & Buchel, C. (2011). Functional dissociation of hippocampal mechanism during implicit learning based on the domain of associations. *Journal of Neuroscience, 31,* 13739–13745.

Rubin, S. R., Van Petten, C., Glisky, E. L., & Newberg, W. N. (1999). Memory conjunction errors in younger and older adults: Event-related potential and neuropsychological data. *Cognitive Neuropsychology, 16,* 459–488.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 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. *Annual Review of Psychology, 70,* 27–52.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Sasaki, Y., & Putz, B. (1998). Transition of brain activation from frontal to parietal areas in visuomotor sequence learning. *Journal of Neuroscience, 18,* 1827–1840.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience, 8,* 657–661.

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, 26,* 1736–1747.

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. *Current Biology, 22,* 1622–1627.

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron, 37,* 1013–1025.

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,* 165–178.

Shohamy, D., & Wagner, A. D. (2008). Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron, 60,* 378–389.

Simon, J. R., Howard, J. H., & Howard, D. V. (2011). Age differences in implicit learning of probabilistic unstructured sequences. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 66,* 32–38.

Simon, J. R., Vaidya, C. J., Howard, J. H., & Howard, D. V. (2012). The effects of aging on the neural basis of implicit associative learning in a probabilistic triplets learning task. *Journal of Cognitive Neuroscience, 24,* 451–463.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences, 13,* 403–409.

Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology-General, 134,* 552–564.