Learning multiple concurrent higher-order visual sequences in human primary visual cortex

Clive R. Rosenthal[1], Indira Mallik[2], Cesar Caballero-Gaudes[3], Marty Sereno[4], and David Soto[3,5]

[1] *Nuffield Department of Clinical Neurosciences, University of Oxford, UK*
[2] *Division of Brain Sciences, Imperial College London, UK*
[3] *Basque Center on Cognition, Brain and Language, San Sebastian, Spain*
[4] *Department of Psychology, San Diego State University, USA*
[5] *Ikerbasque, Basque Foundation for Science, Bilbao, Spain*

January 19, 2018

Correspondence to: d.soto@bcbl.eu, Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, 2nd Floor 20009 San Sebastian

Abstract

Learning and memory are known to be supported by a network involving the medial temporal lobe and linked neocortical regions. Emerging evidence indicates that primary sensory regions (i.e., V1) may be also contribute to recognition memory processes, but this has been only tested with a single visuospatial sequence as the target memorandum . The present study used functional magnetic resonance imaging to test the capacity of human V1 to support the learning of multiple, concurrent and complex visual sequences involving discontinous (second-order) associations. Two peripheral, task-irrelevant but structured sequences of orientated gratings appeared simultaneously in right and left visual fields alongside a central, task-relevant sequence that was in the focus of spatial attention. Pseudorandom sequences were interspersed amongst the three complex structured sequences with different asynchronies. We found that a network involving the precuneus and V1 was involved in learning the structured sequence presented at central fixation, whereas right V1 was modulated by repeated exposure to a concurrent structured sequence presented in the left visual field. The same result was not found in left V1. These results indicate that human V1 can support the learning of multiple concurrent sequences involving complex discontinuous inter-item associations, even those that are goal-irrelevant.

1 Introduction

Primary visual cortex (V1) is typically thought to carry out primitive low-level visual computations and perceptual learning of low-level visual features (Sasaki, Nanez, & Watanabe, 2010), but there is now emerging evidence that V1 can play a role in higher-order functions such as the learning and recognition of visuo-spatial sequences in humans (Rosenthal, Andrews, Antoniades, Kennard, & Soto, 2016), and in model organisms (Cooke, Komorowski, Kaplan, Gavornik, & Bear, 2015; Gavornik & Bear, 2014). Functional coupling between V1 and putative memory substrates (i.e., hippocampus) has been identified during non-conscious recognition (Rosenthal et al., 2016), incidental statistical learning of visible items (Turk-Browne, Scholl, Johnson, & Chun, 2010), conscious recall of tone-grating pairs (Bosch, Jehee, Fernández, & Doeller, 2014) and cue-action associations (Hindy, Ng, & Turk-Browne, 2016). However, the capacity limitations of V1 activity during the learning of complex memoranda remain to be systematically evaluated.

Hence, the aim of the present functional magnetic resonance imaging (fMRI) study was to understand the capacity of human V1 to support concurrent learning and memory when memoranda are presented at attended (task-relevant) locations and unattended (task-irrelevant) peripheral regions of the visual field subtended by V1. Prior research has probed the role of V1 in learning and memory for single or simple visual sequences (Rosenthal et al., 2016; Ekman, Kok, & de Lange, 2017; Cooke et al., 2015; Gavornik & Bear, 2014). Here, we tested the capacity of human V1 for learning multiple sequences based on discontinuous associations of visual events at three different positions in the visual field, both task-relevant and task-irrelevant. Manipulating task-relevance factors is fundamental to defining the role of V1 in higher-order learning. Activity in V1 may reflect the learning computation in question or may be driven by strategic goal-relevant factors, such as top-down attention (Li, Piëch, & Gilbert, 2004; Gandhi, Heeger, & Boynton, 1999; Liu, Larsson, & Carrasco, 2007), which are known to impinge upon V1 responses. Changes in V1 activity associated with the repeated presentation of a task-irrelevant, peripheral unattended but structured stimulus sequence would indicate that learning-related modulation can occur outside of strategic factors related to top-down attentional influences.

We thus assessed the brain responses in V1 associated with the repeated presentation of a visual sequence presented at fixation and also tested whether right and left V1 exhibited learning-related activity associated with the repeated presentation of the peripheral, task-irrelevant, but structured, higher order visual sequences in its corresponding contralateral visual receptive fields. Evidence of V1 modulation in either unattended location would be consistent with a model of V1 at the interface between perception and memory functions, which can operate independently of goal-directed attention control factors.

2 Methods

2.1 Participants

Seventeen undergraduate students were recruited (mean age of 22 years; nine female). All participants gave informed written consent to take part in accordance with the terms of approval granted by the local research ethics committee and the principles expressed in the Declaration of Helsinki, and were naive to the purpose of the experiment. All participants reported normal or corrected-to-normal vision, and had no history of neurological disease.

2.2 Experimental procedure

Participants underwent fMRI scanning as they viewed a visual display containing three different visual sequences. The sequence presented in the center of the screen involved changes in the color of a circle (white, light grey, dark grey and black). The sequences presented in the left and right visual fields involved changes in the orientation of a grating (i.e., 0, 45, 90 or 135 degrees from the vertical) . Each sequence was comprised of 12 stimuli following a second-order conditional (SOC) rule (Reed & Johnson, 1994), whereby at the lowest structural level the current stimulus is contingent on the two preceding stimulus.

Participants were instructed to discriminate the current color of the central sequence by pressing one of four buttons. The central sequence was goal-relevant and hence in the focus of attention. The peripheral sequences were designated task-irrelevant by instructing participant to ignore the changes in orientations of the gratings and only respond to the colour changes in the central sequence and not to respond to the orientations of these peripheral sequences, which were therefore goal-irrelevant. The phase of the gratings was varied independently of the sequences of orientations, and varied randomly across trials to preclude neural adaptation effects of the phase dimension. The three SOC sequences at each of the three locations were presented on a loop, but were asynchronously interspersed with the presentation of the 12 element pseudorandom sequence (i.e., containing no immediate repetitions of a colour or grating orientation). Pseudorandom sequences were always repeated 4 times, hence comprising 48 stimulus presentations or trials. Stimuli at all three locations were each presented for 1 second with an interval of 0.25 seconds between each stimulus. The pseudorandom sequence appeared twice per run at different times for each of the 3 sequences. The temporal position within a run of each of the pseudorandom sequences was consistent across participants. The Table below illustrates the structure of each of the 3 training runs. There were 672 trials on each of the three training runs and each lasted for 14 minutes. The experimental task is depicted in Figure 1.

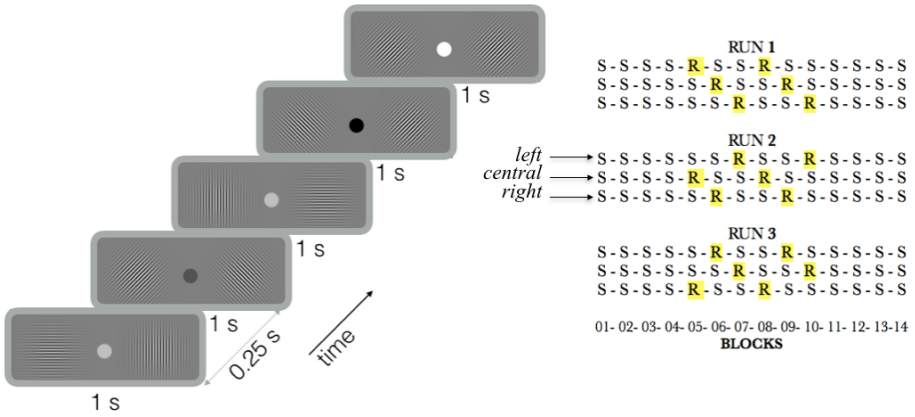

Figure 1: Illustration of the sequence training protocol inside the MRI scanner. The right panel indicates the structure of each of the 3 training runs (S: for the visual second-order conditional sequence; R: for pseudorandom blocks, comprised of a 48 trials each). The top row corresponds to the left sequence, the middle to the central sequence and the bottom row within each run corresponds to the right sequence. Note that each of the pseudorandom sequences were presented asynchronously across the three sequences and scanning runs.

2.3 Recognition test outside of the scanner

After training, the participants performed three recognition tests outside the scanner, one for each of the three 12-element second-order conditional sequences presented at central fixation and left/right positions of the visual field and counterbalanced across participants. On each trial of the recognition test, participants were shown either an old or a new sequence which served as a recognition cue (six-element sequences; 12 old and 12 new). Old sequences were structured sequences that were presented during the training phase inside the scanner and were presented in the same way as during the training protocol. New sequences were also based on a SOC rule, but had not been presented to the participant during the training phase. The participant was asked to indicate whether each of the sequences presented on each trial was 'old' or 'new'. Then, participants indicated the confidence in their response (from 1-guess; 2-fairly confident; 3-sure). There were no restrictions on the participants when viewing the central sequences and hence they were allowed to simulate the motor responses given during training.

Receiver-operating-characteristic (ROC) analyses were used to derive measures of type-1 sensitivity (i.e., ability to distinguish old/new sequences) and type-2 sensitivity (i.e. how confidence relates to memory accuracy). For the type-1 analyses, the 'signal' and 'noise' were defined as old and new retrieval cues, respectively. A 'hit' was, therefore, a correct response ('old') to an old trained (six-element sequence) retrieval cue and a 'correct rejection' was a cor-

rect response ('new') to a new untrained (six-element sequence) retrieval cue. A 'false alarm' was an incorrect response ('old') to a new retrieval cue and a 'miss' was an incorrect response ('new') to an old retrieval. To obtain the type-1 ROC curves, we plotted the probability of hits as a function of the probability of false alarms for all possible decision criteria. The different points in the ROC curve were obtained by calculating cumulative probabilities for hits and false alarms along a recoded confidence continuum ranging from 3 (i.e., certain) to 1 (i.e.. least certain) in signal present trials (old retrieval cues), and from 3 (i.e., certain) to 1 (i.e., least certain) in noise trials (new retrieval cues). On the basis of simple geometry, we computed the area under the ROC curve as a distribution-free measure of the discriminability. For the type-2 sensitivity analyses, the area under the ROC curve was calculated by plotting the cumulative probability of correct responses (either hit or correct rejection) and the cumulative probability of incorrect responses (either false alarm or miss) across the different levels of confidence. The area under the ROC curve was estimated as distribution-free measure of metacognitive ability (Fleming, Weil, Nagy, Dolan, & Rees, 2010; Kornbrot, 2006). Data processing and analyses were performed in R (version 3.2.2).

2.4 MRI data acquisition

MRI scanning was performed in a Siemens Avanto 1.5T MRI scanner using a receive-only 32-channel head coil and body transmit. A screen was mounted at the end of the scanner bore and visual stimuliwere projected onto it from the console room through a wave guide. Subjects viewed the screen using a 45 deg mirror. Functional volumes for both the training phase, functional localiser, and retinotopy consisted of multi-slice T2*-weighted echoplanar images (EPI) with blood oxygenation level dependent (BOLD) contrast with a multiband acceleration factor of 4. We used the following scanning parameters to achieve whole brain coverage: TR = 1000 ms, TE = 54.8 ms, 40 coronal slices, 3.2 mm slice thickness, no interslice gap, and FoV = 205x205 mm (3.2 x 3.2 mm in-plane voxels). There were 844 whole-brain scans per training run. To facilitate anatomical localization and cross-participant alignment, a high-resolution 1x1x1 mm whole-brain structural T1-weighted, magnetization prepared rapid gradient echo (MP-RAGE) scan was acquired for each participant (FOV = 256x224 mm, 176 partitions, TR = 8.4 ms, TE = 3.57 ms, TI = 1000 ms, inversion spacing = 2730 ms, BW = 190 Hz/pixel).

2.5 Functional MRI preprocessing

We used FEAT (fMRI Expert Analysis Tool) Version 6.0, as part of FSL (www.fmrib.ox.ac.uk/fsl). The first 10 EPI volumes were removed to account for T1 equilibrium effects. Non-brain removal was performed using Brain Extraction Tool. Volume realigment of functional scans was carried out using FMRIB's Linear Image Registration Tool MCFLIRT. Scans were realigned relative to the middle scan. We applied a 50 s high-pass temporal filtering to

remove low frequency noise, and spatial smoothing using a FWHM Gaussian kernel of 6 mm.

2.6 Retinotopic mapping

Retinotopic mapping was performed to identify right and left human V1 in each individual. We used standard phase-encoded polar angle retinotopic mapping (Sereno et al., 1995). Participants fixated centrally while viewing a clockwise or counterclockwise rotating pie-shaped 'wedge' with a polar eccentricity. The duration of the scan was 8 min 32 secs (512 volumes). Fourier methods were applied to obtain polar angle maps of the cortex and delineate the borders of right and left V1.

2.7 Functional localizer

A functional localiser scan was conducted in all participants and involved the presentation of items that were flashed at the same three different locations. A block design was used. At each location, there were 24 presentations of gratings (250 ms each with a 250 ms inter-grating interval) in pseudorandom orientations. Following each block, there was a rest period of 12 seconds. The phase of the gratings also varied randomly across trials. We obtained contrast of parameter estimates of brain responses to the spatial position of the peripheral sequences (left>right and right>left positions).

The parameter estimates from the localiser were used to define another mask comprising the voxels in the V1 retinotopic masks that responded to the position of the gratings during the training phase (e.g., define the voxels within the right V1 mask that showed a response to the left-side position of the gratings and likewise for left V1). This area in V1 was delineated by finding the intersection between the V1 retinotopy and the functional localiser maps using fslmaths -mas tool. Hence, we isolated the area of V1 that was most selective of the location of the right and left sequences. We initially derived masks comprising the 10 most responsive voxels and repeated the procedure considering the 5, 20 and 40 most responsive voxels to further assess the reliability of the results.

2.8 MRI Statistical Analyses

Time-series statistical analyses were conducted using FILM (FMRIB's Improved Linear Model) with local autocorrelation correction. The data were analyzed using voxelwise time series analysis within the framework of the general linear model. A design matrix was generated in which each individual presentation of a grating was modelled with a double-gamma hemodynamic response function. Explanatory regressors were created for each of the structured and pseudorandom sequences. Standard motion realignment parameters were included in each individual subject's general linear model as nuisance regressors. We included further nuisance regressors for those volumes corresponding to motion outliers using the FSL motion outliers function. The root mean squared (RSM) head

position difference to the reference volume was used as metric to detect fMRI timepoints corrupted by large motion. The threshold was set to the FSL default, namely, the 75 percentile + 1.5 inter quartile range of the distribution of RSM values of each run. A confound matrix was generated and used in the GLM to completely remove the effects of these timepoints on the analysis. This is intended to deal with the effects of intermediate to large motions, which corrupt images beyond anything that the linear motion parameter regression methods can fix.

Sequence blocks were modeled in chunks of 48 trials in order to contrast activity between structured and pseudorandom baseline blocks (S and R in Figure 1); 48 trials were selected as a multiple of 12-element unit of SOC sequence that allowed the number of trials on structured sequence blocks to be equated with the pseudorandom blocks, as per our prior study (Rosenthal et al., 2016).

For the analysis of the central sequence, we first derived contrast of parameter estimates between the structured relative to the pseudorandom sequences (S<R) within each run. S<R contrast of parameter estimates were obtained as follows. First, given that in some runs (e.g. run 3, see Figure 1) the first central pseudorandom sequence of the block was immediately preceded by additional pseudorandom sequences appearing in the right and left visual fields, we selected a prior central structured block that (i) was not flanked by pseudorandom sequences on the left and right. Therefore, the two structured blocks preceding the first central pseudorandom were not considered in the S<R contrast. Instead, we selected the preceding structured block (i.e., the third one preceding the first central pseudorandom sequence, see Figure 1). This was done to avoid potential interference from peripheral pseudorandom sequences presented concurrently with the critical central structured sequence. Secondly, we performed the same analyses by comparing the structured sequence block immediately preceding a pseudorandom sequence block at similar timepoint across the runs. The results obtained from these two analysis strategies were similar.

Following the computation of within-run contrast of parameter estimates (S<R), we performed across-run within-subject (fixed-effects) analyses using lower level parameter estimates to test for a linear training effects across the 3 runs (e.g., -1 0 1) and likewise exponential training effects (e.g., -2 -0.5 2.5). Finally, the parameter estimates from these linear and exponential contrasts across the three learning runs were converted to standard MNI 2 mm template with linear transformation and passed to a higher-level across subjects one-sample t-test to assess brain regions consistent across participants that were associated with increased training effects in the structured relative to the pseudorandom sequence using FLAME (Local Analysis of Mixed Effect). Statistical maps were thresholded using clusters determined by $Z > 2.3$ and a corrected cluster extent significance threshold of $p = 0.05$, using Gaussian Random Field Theory. Across participants one-sample t-tests were also performed separately on each of the 3 runs to assess $S < R$ brain actitivy differences on each run. These considerations applied also to the psychophysiological interaction (PPI) analyses described below.

We also used non-parametric permutation-based one-sample t-tests to fur-

ther assess the statistical significance of the results. The FSL Randomise program was used (Winkler, Ridgway, Webster, Smith, & Nichols, 2014) with threshold-free cluster enhancement (TFCE). TFCE is a method that enhances values within cluster-like structures, while preserving voxel-level inference (Smith & Nichols, 2009). We performd 6 mm spatial smoothing of the variance which is typically included to increase study power.

For the analysis of the peripheral right and left sequence, we adopted the following approach. Parameter estimates were derived in native functional space for each structured and pseudorandom sequence block (Figure 1) for each training run and participant. We then used the V1 masks obtained from the retinotopy and the functional localiser to extract the parameter estimates of V1 activity across the training phase, separately for the right and left V1 masks. For each run, we then performed paired-sample t-tests comparing V1 activity estimates in the first pseudorandom sequence and the inmediately preceding structured sequence.

2.9 Psychophysiological interaction (PPI) analyses

A seed-voxel PPI based approach was used to identify signals of interest that would be missed in standard subtraction based analyses. A 6 mm radius posterior cingulate/precuneus mask was drawn and based on the peak voxels of the posterior cingulate/precuneus that showed effects of training. This mask was used to define the seed region's time course for the PPI analyses. The aim was to assess whether the temporal correlation between posterior precuneus and primary visual cortex was modulated during the learning of the task-relevant, second-order conditional sequence.

We then estimated a model for each participant that included the same psychological regressors as outlined above, for the onsets of each sequence, and, critically, for the PPI model, a physiological regressor for the time course of the region-of-interest, and psychological x physiological interaction regressors for the PPI. These new regressors were, therefore, added to the previous first-level model for each participant/run. Parameter estimates for structured versus pseudorandom conditions based on the PPI regressors were derived in a similar fashion to the analyses outlined above (i.e., using both fixed-effects analysis across runs followed by mixed-effect analysis across participants), which, here, directly compared the changes in functional coupling associated with training. Given our a priori interests in V1(Rosenthal et al., 2016), the higher-level analyses (across subjects) were performed using a region-of-interest approach, with a target occipital mask in standard space derived from the FSL Harvard-Oxford Atlas.

3 Results

3.1 Behavioural performance during training

We assessed the response time (RT) and the accuracy of manual responses to each item color of the central sequence as a function of the nature of the sequence (structured vs pseudorandom). Overall, the results indicate that the repeated exposure to the deterministic discontinuous associations of the structured sequences affected somatomotor performance relative to each novel pseduorandom sequence. In run 1, accuracy on the first block of pseudorandom sequences was lower than on the preceding structured block (mean S=0.873, mean R=0.765; $t(16) = 4.675$, $p = 0.0002$; Student's t-tests presented in this paper are two-tailed). Mean RTs were likewise slower during pseudorandom relative to the preceding structure sequence (mean S=528 ms, mean R=552 ms; $t(16) = 3.219$, $p = 0.0053$). In run 2, the difference in accuracy between the first pseudorandom sequence and the previous structured sequence was also significant (mean S=0.85, mean R=0.789; $t(16) = 2.398$, $p = 0.02899$) and a consonant pattern was found in the RTs (mean S= 514 ms, mean R=540 ms; $t(16) = 2.731$, $p = 0.015$). Finally, in run 3, performance was slower with the onset of the first pseudorandom block relative to the preceding structured sequence (mean S=502 ms, mean R=540 ms; $t(16) = 2.2563$, $p = 0.038$). There were no significant differences in accuracy between the first pseudorandom sequence and the previous structured sequence (mean S=0.827, mean R=0.794; $t(16) = 1.037$, $p = 0.315$). Taken together, these results indicate that the statistical structure of the attended sequence was learnt.

3.2 Recognition performance outside the scanner

Following the fMRI training task, a surprise recognition test was administered to test the memory of participants for each of the 3 sequences. To summarise the main results, participants memory confidence ratings discriminated whether the central (goal-relevant) sequence retrieval cues were 'old' (repeateadly studied during the training runs) or 'new', but there was no evidence of significant recognition memory for the unattended left and right sequences.

The results for the recognition analyses for the sequence presented at central fixation showed that the area under the type-1 ROC for 'old'/'new' discrimination was only marginally above chance ($t(16) = 1.882$, $p = 0.078$, mean = 0.565). Type- 2 ROC performance, namely, the extent to which confidence ratings predicted the accuracy of the recognition response showed a similar trend, but did not surpass statistical significance ($t(16) = 1.911$, $p = 0.074$, mean = 0.544. Notably, participants were more confident in their responses when judging old sequences relative to new sequences ($t(16) = 2.2706$, $p = 0.036$, mean confidence old = 1.926; mean new = 1.784), indicating that confidence was diagnostic of the memory status (old/new) of the sequences.

Further results presented next indicated that participants were unable to recognise either the left or the right sequences.

Analyses of the recognition test performance for the left sequence showed that the area under the type-1 and type-2 ROC was no different from chance ($t(16) = 0.543$, $p = 0.594$, mean = 0.484 for type-1 ROC, and $t(16) = 0.352$, $p = 0.729$, mean = 0.492 for type-2 ROC performance). Moreover, the results showed that participants were not more confident in responding to old vs. new left sequences ($t(16) = 1.0371$, $p = 0.31$, mean confidence old = 1.593; mean new = 1.647).

There was also no evidence of recognition memory for the right sequence. Type-1 and type-2 performance were no different from chance ($t(16) = 1.075$, $p = 0.298$, mean 0.483, and $t(16) = 0.101$, $p = 0.92$, mean = 0.498). Participants confidence ratings did not discriminate between old and new left sequences ($t(16) = 1.165$, $p = 0.261$, mean confidence old = 1.598; mean new = 1.544).

3.3 fMRI Results

A key question of the present study is whether the V1 is involved in supporting the learning of complex deterministic discountinous associations in the structured peripheral sequences of oriented gratings (unattended, task-irrelevant) that participants were unable to recognise in a subsequent test. The behavioural results showed that participants could only learn and recognise the central goal-relevant sequence of color changes.

For the analysis of the central sequence, we computed statistical contrasts of parameter estimates between structured (S) and pseudorandom (R) sequences (i.e., comparing S<R) on each of the three training runs (see Methods). The analysis strategy to compare S vs R sequences across training was as follows. Given that in some runs (e.g., run 3, see Figure 1) the first central pseudorandom (R) block was inmediately preceded by additional pseudorandom sequences appearing in the right and left visual fields, we selected a preceding central structured (S) block that (i) was not flanked by pseudorandom sequences blocks on the left and right. Therefore, the two structured blocks preceding the first central pseudorandom were not considered in this contrast, and instead we selected the third structured block preceding the first central pseudorandom sequence (see Figure 1). This was done to avoid potential interference from peripheral pseudorandom sequences presented concurrently with the critical central structured sequence. We then tested for linear and exponential effects of training on BOLD responses (e.g., increased differential S<R response in run 3 relative to run 1). Learning the structure of the repeated sequence ought to be reflected in brain activity changes across the training runs whenever a non-structured, pseudorandom sequence was presented.

The analyses showed significant clusters in the posterior cingulate gyrus (MNI 0 -48 16, Z=4.26) extending into the posterior precuneus and the vicinity of intracalcarine sulcus, the ventromedial prefrontal cortex (MNI, 2 64 2, Z=3.41). All of these regions displayed significant linear and exponential changes in activity across runs ($Z > 2.3$. $p < 0.05$ whole brain corrected using cluster-based random-field theory with family wise correction).

Secondly, we performed the same analyses by comparing the structured sequence block immediately preceding a pseudorandom sequence block at similar timepoint across the runs. The results obtained from this second analysis was similar to the first except that the ventromedial prefrontal cortex did not survive the threshold for statistical significance. Activity was restricted to the precuneus.

To examine further the effects of training, we performed group-analyses separately for each of the three runs. On the basis of prior studies of learning SOC sequences (Schendan, Searl, Melrose, & Stern, 2003; Albouy et al., 2008), training effects were predicted to be most evident on the later third training run. Differences between the structured and the pseudorandom sequence blocks in run 1 and in run 2 did not survive the above significance threshold. Significant differences in the precuneus and ventromedial PFC emerged only in run 3, and were evident in both linear and exponetial learning contrasts ($p < 0.05$ whole brain corrected and also following non-parametric tests using threshold-free cluster enhacement).

Following prior reports that putative mnemonic regions (e.g. hippocampus) show increased temporal correlation with V1 during memory recall (Bosch et al., 2014)and implicit statistical learning of visible stimuli (Turk-Browne et al., 2010), we conducted similar analyses considering the posterior precuneus cluster that was observed during learning, as this region is also a key node of the memory network (Ranganath & Ritchey, 2012). This region has connections to higher level visual areas (Margulies et al., 2009) and thus has access to V1 via feedback projections.

The psychophysiological interaction (PPI) analysis used a 6 mm mask centered on the posterior precuneus (MNI 45 39 42) at the cluster peak in the S>R contrast in run 3 reported above. This was used to define the seed region's time course for the PPI. Given our a priori interest in primary visual cortex, training effects across participants were assessed within a standardized occipital mask as target of the PPI.

We found a cluster of activity in V1 peaking at the intracalcarine sulcus (MNI: 8 -92 0, Z=3.34) showing linear and exponential increases in temporal correlation with posterior precuneus across the training runs on structured relative to the pseudorandom sequences (see figure 2). This result was observed in the comparison between first pseudorandom sequence block and the structure sequence immediately preceding it. This suggests that learning a representation of the structured sequence involves the coupling between putative memory substrates and V1. For the position of anatomically defined V1 in the calcarine, see (Horton & Hoyt, 1991); also, the activity in the intracalcarine cortex and its vicinity coincides with probabilistic V1.

Right and left V1 were mapped using independent retinotopy and functional localiser scans of each participant (see Methods). It was predicted that learning of the structured visual sequence displayed on the left visual field would be mainly reflected in right V1 responses, whereas the opposite was predicted for the right sequence. To test these predictions, we compared V1 responses to first pseudorandom sequence and the immediately preceding structured sequence, on

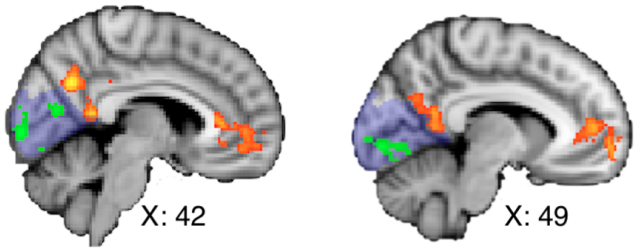

Figure 2: Brain regions in the precuneus and ventromedial prefrontal cortex exhibited linear changes in BOLD signal for the structured<pseudorandom sequence contrast across the three training runs (in yellow, $Z > 2.3$, $p < 0.05$, whole-brain corrected). The green areas in primary visual cortex indicate the results of the PPI analysis from the posterior precuneus cluster ($Z > 2.3$, $p < 0.05$, corrected for the occipital mask in blue shading).

each of the three training runs. If V1 activity is selective for the SOC sequence, this ought to be reflected in brain activity changes across the training runs whenever a non-structured, pseudorandom sequence was presented. Specifically, as with the central sequence, learning related activity in V1 should be most evident in the last phase of training, namely in run 3.

For right V1, in run 3, the BOLD signal difference between structured and pseudorandom sequences was significant ($t(16) = 2.480$, $p = 0.0246$, two-tailed paired t-test). We did not find any significant difference in this contrast during run 1 ($t(16) = 1.028$, $p = 0.318$) and run 2 ($t(16) = 0.239$, $p = 0.813$). These results were consistent across a wide range of V1 masks. In particular, we initially tested our hypothesis with masks that comprised the 10 most responsive voxels to the location of the left sequence, as shown by the combination of retinotopy and the functional localiser, and then verified that the result held with masks containing 5, 20 and 40 most responsive voxels. Figure 3 illustrates the representative results from the 10 most responsive voxels.

For left V1, there was no evidence of learning the structure of the deterministic discontinuous associations of the structured sequence in any of the three runs (run 1: $t(16) = 0.0007$, $p = 0.99$; run 2: $t(16) = 0.112$, $p = 0.912$; run 3: $t(16) = 0.66$, $p = 0.518$). Figure 4 illustrates these results.

4 Discussion

The results indicate that V1 operates as part of a network involving the precuneus during the learning of task-relevant visual sequences involving discontinuous associations and which are within the focus of spatial attention. Precuneus, posterior cingulate and ventromedial prefrontal cortex have been identified as part of a memory network (Ranganath & Ritchey, 2012) which can be

Figure 3: Right V1 responses as a function of the structured (S) or pseudorandom (R) status of the sequences on each of the three training runs. There were significant BOLD signal change between structured and pseudorandom sequences in run 3, but not in run 2 or 1 (see text).

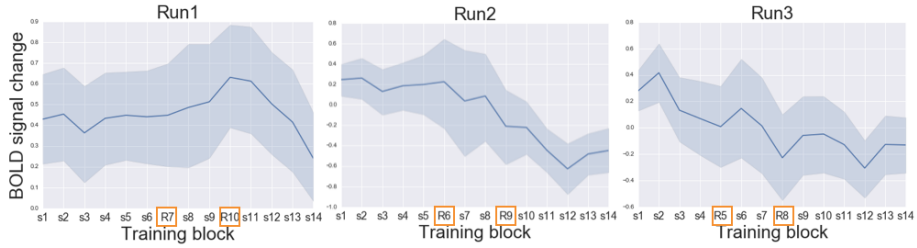

Figure 4: Left V1 responses as a function of the structured (S) or pseudorandom (R) status of the sequences on each of the three training runs. There was no evidence of BOLD signal differences between structured and pseudorandom sequences in any of the runs (see text).

modulated by the temporal structure of relevant memoranda (Hasson, Chen, & Honey, 2015), with higher-order regions exhibiting sensitivity to increasingly longer temporal structures, ranging from tens of seconds (e.g. written sentences) in more posterior areas to minutes (e.g., a movie). V1 coupling with posterior precuneus during repeated exposure to the structured sequences is consonant with a mnemonic feedback signal from higher-order brain regions involved in predictive processing (Rao & Ballard, 1999; Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010).

The second main result revealed that retinotopic V1 areas for non-relevant, non-attended spatial positions appear to support learning the statistical regularities of visual sequences presented concurrently with goal-relevant, attended sequences. Notably, activity in right V1 was associated with learning the SOC sequence presented in the left visual field for which right V1 exhibited the greater spatial selectivity. However, there was no evidence of analogous effects in left V1. It is possible that visual competition amongst the three concurrent sequences reduced overall the capacity of left and right V1 areas to encode the multiple co-ocurring sequences, but there is no immediate reason to explain the

preferential capacity of right V1 activity to support sequence-specific learning alongside a null sensitivity in left V1. Prior work has suggested the existence of lateralised perceptual biases (Railo, Tallus, & Hamalainen, 2011; Nicholls, Mattingley, & Bradshaw, 2005) in which perceptual processing is biased towards the left side of space. If such biases were operating here that could explain the preferential engagement of right V1 for processing the left sequence.

These results are in keeping with recent reports that V1 is subject to experience-dependent changes that support higher-order mnemonic processes (Rosenthal et al., 2016; Ekman et al., 2017; Cooke et al., 2015; Gavornik & Bear, 2014). Our findings further indicate that V1 processing capacity is flexible enough to be deployed even when attention is constrained by visual competition and task-relevance factors. Top-down processes related to attention (Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999)and predictive processing (Kok, Jehee, & De Lange, 2012) are known to modulate V1. However V1 seems capable of operating autonomously during higher-order processing, so that complex knowledge at goal-irrelevant peripheral locations can nevertheless be encoded. This is consistent with a recent observation that V1 can automatically 'preplay' an expected sequence of events, even when attention resources are constrained (Ekman et al., 2017).

Finally, it is worth noting that the mnemonic feedback signals we have observed in V1 may come from different laminae than feedforward signals; experiments at higher resolution will be needed to make this determination.

5 Acknowledgements

D.S. acknowledges support from the Spanish Ministry of Economy and Competitiveness (MINECO), through the 'Severo Ochoa' Programme for Centres/Units of Excellence in R&D (SEV-2015-490) and project grants PSI2016-76443-P from MINECO and PI-2017-25 from the Basque Government.

References

Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., ... others (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron*, *58*(2), 261–272.

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, *30*(8), 2960–2966.

Bosch, S. E., Jehee, J. F., Fernández, G., & Doeller, C. F. (2014). Reinstatement of associative memories in early visual cortex is signaled by the hippocampus. *Journal of Neuroscience*, *34*(22), 7493–7500.

Cooke, S. F., Komorowski, R. W., Kaplan, E. S., Gavornik, J. P., & Bear, M. F. (2015). Visual recognition memory, manifested as long-term habituation, requires synaptic plasticity in v1. *Nature neuroscience*, *18*(2), 262–271.

Ekman, M., Kok, P., & de Lange, F. P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. *Nature Communications*, *8*.

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), 1541–1543.

Gandhi, S. P., Heeger, D. J., & Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences*, *96*(6), 3314–3319.

Gavornik, J. P., & Bear, M. F. (2014). Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nature neuroscience*, *17*(5), 732–737.

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in cognitive sciences*, *19*(6), 304–313.

Hindy, N. C., Ng, F. Y., & Turk-Browne, N. B. (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature neuroscience*, *19*(5), 665.

Horton, J. C., & Hoyt, W. F. (1991). The representation of the visual field in human striate cortex: a revision of the classic holmes map. *Archives of ophthalmology*, *109*(6), 816–824.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*(4), 751–761.

Kok, P., Jehee, J. F., & De Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–270.

Kornbrot, D. E. (2006). Signal detection theory, the approach of choice: Model-based and distribution-free measures and evaluation. *Attention, Perception, & Psychophysics*, *68*(3), 393–414.

Li, W., Piëch, V., & Gilbert, C. D. (2004). Perceptual learning and top-down influences in primary visual cortex. *Nature neuroscience*, *7*(6), 651.

Liu, T., Larsson, J., & Carrasco, M. (2007). Feature-based attention modulates orientation-selective responses in human visual cortex. *Neuron*, *55*(2), 313–323.

Margulies, D. S., Vincent, J. L., Kelly, C., Lohmann, G., Uddin, L. Q., Biswal, B. B., ... Petrides, M. (2009). Precuneus shares intrinsic functional architecture in humans and monkeys. *Proceedings of the National Academy of Sciences*, *106*(47), 20069–20074.

Nicholls, M. E., Mattingley, J. B., & Bradshaw, J. L. (2005, September). The effect of strategy on pseudoneglect for luminance judgements. *Cognitive Brain Research*, *25*(1), 71–77. Retrieved 2017-10-07, from `http://linkinghub.elsevier.com/retrieve/pii/S092664100500114X` doi: 10.1016/j.cogbrainres.2005.04.006

Railo, H., Tallus, J., & Hamalainen, H. (2011). Right visual field advantage for perceived contrast: Correlation with an auditory bias and handedness.

Brain and Cognition, *77*(3), 391–400.

Ranganath, C., & Ritchey, M. (2012). Two cortical systems for memory-guided behaviour. *Nature reviews. Neuroscience*, *13*(10), 713.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, *2*(1).

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(3), 585–594. Retrieved 2017-12-04, from `http://doi.apa.org/getdoi.cfm?doi=10.1037/0278-7393.20.3.585` doi: 10.1037/0278-7393.20.3.585

Rosenthal, C. R., Andrews, S. K., Antoniades, C. A., Kennard, C., & Soto, D. (2016). Learning and recognition of a non-conscious sequence of events in human primary visual cortex. *Current Biology*, *26*(6), 834–841.

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature reviews. Neuroscience*, *11*(1), 53.

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003, March). An fMRI Study of the Role of the Medial Temporal Lobe in Implicit and Explicit Sequence Learning. *Neuron*, *37*(6), 1013–1025. Retrieved 2017-12-04, from `http://linkinghub.elsevier.com/retrieve/pii/S0896627303001235` doi: 10.1016/S0896-6273(03)00123-5

Sereno, M., Dale, A., Reppas, J., Kwong, K., Belliveau, J., Brady, T., & Rosen BR, R., Tootell. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–93.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, *44*(1), 83–98.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience*, *30*(33), 11177–11187.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *Neuroimage*, *92*, 381–397.