

# Distractor-resistant Short-Term Memory Is Supported by Transient Changes in Neural Stimulus Representations

Jan Derrfuss<sup>1,2</sup>, Matthias Ekman<sup>1</sup>, Michael Hanke<sup>3,4</sup>,  
Marc Tittgemeyer<sup>5</sup>, and Christian J. Fiebach<sup>1,6,7</sup>

## Abstract

■ Goal-directed behavior in a complex world requires the maintenance of goal-relevant information despite multiple sources of distraction. However, the brain mechanisms underlying distractor-resistant working or short-term memory (STM) are not fully understood. Although early single-unit recordings in monkeys and fMRI studies in humans pointed to an involvement of lateral prefrontal cortices, more recent studies highlighted the importance of posterior cortices for the active maintenance of visual information also in the presence of distraction. Here, we used a delayed match-to-sample task and multivariate searchlight analyses of fMRI data to investigate STM maintenance across three extended delay phases. Participants maintained two samples (either faces or houses) across an unfilled pre-distractor delay, a distractor-filled delay, and an unfilled post-distractor delay. STM contents (faces vs. houses) could

be decoded above-chance in all three delay phases from occipital, temporal, and posterior parietal areas. Classifiers trained to distinguish face versus house maintenance successfully generalized from pre- to post-distractor delays and vice versa, but not to the distractor delay period. Furthermore, classifier performance in all delay phases was correlated with behavioral performance in house, but not face, trials. Our results demonstrate the involvement of distributed posterior, but not lateral prefrontal, cortices in active maintenance during and after distraction. They also show that the neural code underlying STM maintenance is transiently changed in the presence of distractors and reinstated after distraction. The correlation with behavior suggests that active STM maintenance is particularly relevant in house trials, whereas face trials might rely more strongly on contributions from long-term memory. ■

## INTRODUCTION

Short-term memory is the ability to actively maintain task-relevant information over brief periods of time. Monkeys and humans can maintain such information even when distractors intervene between encoding and recall. In a now classic study, Miller, Erickson, and Desimone (1996) showed that individual lateral prefrontal neurons maintained sample-selective delay activity in a delayed match-to-sample (DMS) task despite multiple distractors intervening between sample and probe, with sample-selective delay activity defined as increased delay-related activity to the “preferred” stimulus of the respective neuron. Recording from prefrontal neurons that showed a modulation of activity in at least one phase of their DMS task, Miller et al. (1996) showed sample-selective delay activity in 28% of these neurons. This finding led them to conclude that the lateral pFC plays an important role in distractor-resistant STM.

Interestingly, however, it seems that neurons with sample-selective delay activity did not show sample selectivity while the distractors were presented (see Figure 5 in Miller et al., 1996). Thus, the question arises—how is information about the sample maintained during distractor presentation—a question that was not addressed by Miller and colleagues. We hypothesize that, in principle, there are two possible mechanisms by which sample-related information may survive distractor delays: active and passive maintenance. Active maintenance would involve persistent neuronal activity, although the level of neuronal activity might be substantially reduced relative to stimulus presentation. As this form of maintenance requires ongoing neuronal activity, its metabolic effects should in principle be detectable with fMRI. Passive maintenance, on the other hand, could be described as a state of heightened accessibility of information (as, e.g., assumed in the working memory model of Cowan, 2001), in the absence of active maintenance and thus also without persistent neuronal firing or increased fMRI activation.

Only a limited number of studies have so far investigated short-term or working memory maintenance in the presence of distractors. In one of these studies, Jacob and Nieder (2014) used a delayed match-to-numerosity

<sup>1</sup>Radboud University Nijmegen, <sup>2</sup>University of Nottingham, <sup>3</sup>Otto von Guericke University, Magdeburg, Germany, <sup>4</sup>Center for Behavioral Brain Sciences, Magdeburg, Germany, <sup>5</sup>Max Planck Institute for Metabolism Research, Cologne, Germany, <sup>6</sup>Goethe University Frankfurt, <sup>7</sup>Center for Individual Development and Adaptive Education, Frankfurt am Main, Germany

paradigm to study distractor-resistant STM in macaque monkeys while recording simultaneously from lateral pFC and the ventral intraparietal area (VIP). Monkeys had to maintain information about the number of dots shown as sample across a pre-distractor delay, a distractor delay, and a post-distractor delay; distractors were also dots, but their numerosity was task-irrelevant. In line with Miller et al. (1996), Jacob and Nieder found that prefrontal neurons displayed sample-selective delay activity even after distraction. Going beyond Miller et al., Jacob and Nieder also studied neuronal activity during the distractor delay itself and found that distractors interfered with numerosity representations in pFC, suggesting that STM representations may not be actively maintained in the pFC while distractors are presented. Unexpectedly, however, Jacob and Nieder found that VIP neurons displayed distractor-resistant response properties: Many VIP neurons maintained sample selectivity even in the distractor delay. These findings suggest that posterior, not lateral prefrontal, cortices play a major role in representing distractor-resistant information in STM.

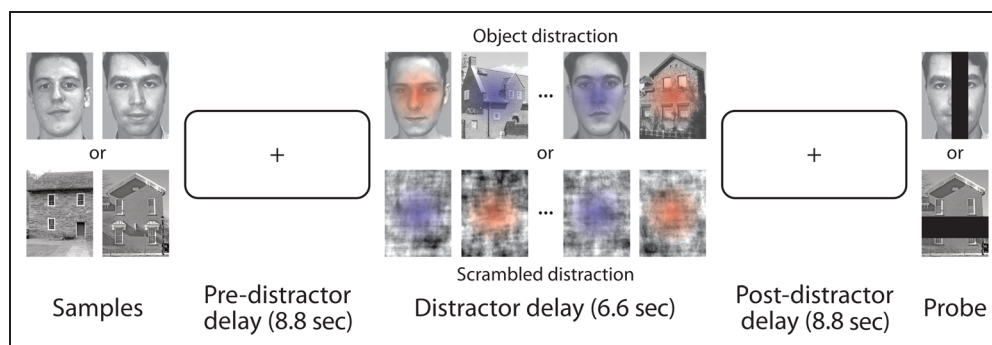
Initial support for this hypothesis for humans comes from an fMRI DMS study with faces and shoes as samples and distractors ( $2 \times 2$  factorial design), reported by Jha, Fabian, and Aguirre (2004). These authors observed increased activity in the right fusiform face area (FFA) when participants maintained faces in memory and faces were at the same time presented as distractors. This could suggest that FFA was involved both in maintaining the memoranda and processing the distractors. However, more detailed analyses did not fully support this assumption, leading the authors to conclude that heightened FFA activity might in fact be due to increased FFA inhibition to filter out distractors when samples and distractors come from the same category, or due to an interaction of maintenance and inhibition. More recently, Bettencourt and Xu (2016) investigated the roles of visual areas V1 to V4 and of a load-sensitive segment of the posterior intraparietal sulcus in distractor-resistant STM. They demonstrated that the orientation of grating stimuli maintained in memory could be decoded from the posterior intraparietal sulcus during the presence of distraction (DMS task; delay length = 11 sec). In addition, when the presence of distractors was not predictable, distractor-resistant maintenance was also found in lower visual areas V1–V4. These findings, thus, provide converging evidence that, also in humans, posterior cortices are involved in the distractor-resistant maintenance of information in STM.

However, although the studies of Bettencourt and Xu (2016) and Jha et al. (2004) suggested that—at least under some conditions—there is active maintenance of information during distractor delays, this has not been a consistent finding. Using a variant of the DMS task, Lewis-Peacock, Drysdale, Oberauer, and Postle (2012; see also Lewis-Peacock & Postle, 2012) found that the sample category could not be decoded once the distrac-

tor presentation commenced, despite successful performance on the subsequent probe presentation. Instead, these authors were able to decode the category of the distractor pictures, even though these were irrelevant to the task. In Lewis-Peacock et al.'s (2012) Experiment 2, two samples from two different stimulus categories were encoded. Participants initially maintained both samples (Delay 1: 8 sec), then were cued to attend to one of the samples (Delay 2: 7.5 sec), before being cued to continue attending to the same sample or to switch their attention to the other sample (Delay 3: 8 sec). During Delay 2, only the category of the attended sample could be decoded, even though the unattended sample might again become relevant again in Delay 3 and in that case also could be decoded during Delay 3. The results of Lewis-Peacock et al. suggest that successful performance in STM tasks does not always have to rely on active maintenance of sample information; in tasks involving diversion of attention away from STM contents, it seems to rely partly on what we have described above as passive maintenance and a subsequent reactivation or recovery stage once the distraction ends (cf. Sprague, Ester, & Serences, 2016). As a result of this, the sample information would then again be held in STM (and thus become accessible to fMRI-based decoding). There is some evidence that such recovery after distraction is supported by medial-temporal lobe structures (cf. Sakai & Passingham, 2004; Sakai, Rowe, & Passingham, 2002a).

Previously, a number of further influential fMRI studies have investigated the effects of distractors presented in the encoding phase or the effects of briefly presented distractors in the delay phase (e.g., Bloemendaal et al., 2015; Clapp, Rubens, & Gazzaley, 2010; Zanto & Gazzaley, 2009; Yoon, Curtis, & D'Esposito, 2006; Gazzaley, Cooney, Rissman, & D'Esposito, 2005). However, as a result of their design, these studies provide no further evidence regarding whether or not STM contents are actively maintained in the distractor delay and, if so, whether or not the maintenance-related neural processes in the distractor delay differ from those in unfilled delays. To investigate these questions, the present fMRI study utilized a variant of the DMS task in which participants were asked to maintain two faces or two houses over three extended delay periods, with distractors being presented in the middle delay period (Figure 1). This design allowed us to apply multivariate pattern analysis (MVPA) to (i) decode which type of stimulus (faces vs. houses) is maintained in memory within each individual delay period (within-delay analyses) as well as to (ii) investigate the generalization of activity patterns across delay periods (across-delay analyses). By presenting pictures of faces and houses as well as scrambled pictures as distractors, we furthermore investigated (iii) how the perceptual similarity of distractors and samples affected the decodability of sample information. In addition to the DMS task, we acquired an independent data set to functionally localize face-preferential and house-preferential

**Figure 1.** The DMS task employed in this study. Participants had to encode and maintain two sample pictures (faces or houses) across three delay periods. The first delay period was unfilled (pre-distractor delay). For both face and house trials, the middle delay period was filled with distractor pictures, which were either intact or scrambled (distractor delay). Overlaid on the distractors was a slight blue or red color gradient (exaggerated in the figure to increase visibility) to which participants responded with index or middle finger button presses, respectively. The final delay period was again unfilled (post-distractor delay). At the end of the trial, a probe picture was presented that matched one of the sample pictures with a probability of 50%. About 25% of the probe stimulus was unpredictably covered to discourage the reliance on specific, salient features during stimulus encoding. All pictures were presented individually in the center of the screen.



brain regions to investigate if these stimulus-preferential areas (in particular FFA and the parahippocampal place area [PPA]) overlap with areas exhibiting above-chance decoding in the DMS task.

## METHODS

### Participants

Twenty-two participants (10 women) took part in the experiment. One male participant was excluded because of multiple movements clearly exceeding the size of one voxel in the functional runs. The mean age of the remaining 21 participants was 27.1 years ( $SD = 3.3$ , range = 22–34 years). All participants were right-handed (laterality quotient of  $>50$  in the Edinburgh Inventory; Oldfield, 1971), had normal or corrected-to-normal vision, and reported no deficits in color vision. No participant reported a history of neurological, major medical, or psychiatric disorder. The study was approved by the local ethics committee, and written consent was obtained from all participants.

### Design and Stimuli of the DMS Task

This study employed a variant of the DMS task (Figure 1). On each trial, participants encoded either two faces or two houses (2.1 sec each, intertrial interval = 0.1 sec) and maintained these stimuli across three delay periods. The first delay was unfilled (fixation cross, 8.8 sec). The second delay was filled with six distractor pictures (6.6 sec, 1.1 sec per distractor): In the object condition, pictures of three faces and three houses were sequentially presented as distractors (distractors came from both categories to keep the distractor-related visual input for faces and houses constant in all trials). In the scrambled condition, six phase-scrambled pictures were presented. To ensure that distractors were attended, all distractor pictures were overlaid with a slight blue or red color gradient, and participants had to indicate the color by a button press. Responses were given with the right index finger (blue) or the right middle finger (red); maximum

RT was 1.1 sec. The order of colors and distractor categories was randomized with the constraint that the same color or distractor category could not appear more than twice in a row. The distractor delay was followed by a final delay phase that was again unfilled (fixation cross, 8.8 sec). Next, a probe picture was presented. The probability of a match to one of the samples was 50%, equated over the two samples. Responses were again given with the right index finger (match) or the right middle finger (no-match; maximum RT = 3 sec). About 25% of the probe stimulus were unpredictably covered, either in vertical or horizontal direction (cf. Figure 1), to ensure that participants could not successfully perform the task by simply encoding a specific feature of the stimulus (e.g., an unusually shaped mouth) and compare this feature with the probe stimulus. Overall, there were four conditions: face samples/object distraction, face samples/scrambled distraction, house samples/object distraction, and house samples/scrambled distraction. Each trial was followed by intertrial intervals of varying lengths (4.4, 6.6, or 8.8 sec), resulting in overall trial lengths of 35.2, 37.4, or 39.6 sec (uniform distribution, controlled on a per-condition basis).

The task was presented in two runs. Each run lasted 24.2 min and began with four dummy scans (to achieve steady-state magnetization), followed by a rest phase of 44 sec (used for echo weighting, see below). Each run involved three task blocks of 7.5 min each. After each task block, participants received feedback about their accuracy (separately for the memory probes and the color decision task; 2 sec) and rested for 17.6 sec. Each block consisted of 12 trials (three per condition). Accordingly, 72 trials were presented overall (18 per condition). In each block, direct repetitions of conditions were excluded, and the remaining transitions were counterbalanced. Each block had six match and six no-match trials. Moreover, after two blocks the number of match and no-match trials for each condition was equalized (as there were three trials per condition per block, this was not possible within a single block).

All pictures were presented individually in grayscale in the center of the screen and subtended a visual angle of  $2.6 \times 3.5^\circ$ . Pictures were shown only once as samples (thus, every encoding period involved two pictures that were not seen before). However, the no-match probes were previously presented as sample stimuli to make sure that participants could not base their decision solely on stimulus familiarity (i.e., on whether or not they had seen a particular face or house before). Overall, 72 face and 72 house pictures were presented as samples and probes (if the very first trial of the experiment was a no-match trial, one additional probe picture was presented). The face pictures came from a set assembled by Endl et al. (1998), depicting men photographed in front of a uniform gray background (see Figure 1). House pictures came from a set of face and house pictures assembled by Piekema, Kessels, Rijpkema, and Fernandez (2009). The left and right edges of house pictures were cropped to achieve the same aspect ratio as that of the face pictures (see Figure 1). We only chose pictures in which the house was still clearly visible after cropping.

Distractors also came from the set assembled by Piekema et al. (2009) but were not used as samples or probes. Distractors were shown twice, with a minimum of 36 intervening distractors before a repetition of the same picture could occur. To generate scrambled distractors, MATLAB version 2014b (The MathWorks, Natick, MA) was used. For each picture, a fast Fourier transform was performed. Then, random phase information was added and an inverse fast Fourier transform was performed to generate the scrambled picture. The color gradient overlaid onto the distractors was Gaussian-shaped with the highest color intensity located in the middle of the image. Overall, 54 face and 54 house distractors were used.

### Design and Stimuli of the Functional Localizer Task

In the functional localizer task, participants had to decide for each presented stimulus whether or not it matched the immediately preceding stimulus (1-back task; right index finger: match; right middle finger: nonmatch). There were separate blocks for face and house stimuli, and 10 trials were presented per block. Half of the blocks of each stimulus category had one 1-back target and the other half had two 1-back targets, to ensure that participants would maintain attention up to the end of the block. Participants were instructed not to respond to the first stimulus in each block (as a 1-back decision is not possible). The 1-back run lasted about 7 min and began with four dummy scans (to achieve steady-state magnetization), followed by a rest phase of 44 sec (used for echo weighting, see below). Participants performed six face and six house blocks in alternating order; 50% starting with a face block and the other half with a house block. A block lasted for 15 sec and was followed by a 15-sec rest period. Pictures were presented for 1 sec, followed by a blank screen of 500 msec. The pictures

used were a randomly selected subset of the pictures used as samples in the DMS task. No pictures apart from the 1-back targets were repeated.

### Procedures

After being welcomed, participants were given general information about the study, received an MRI participant information form, and gave informed consent. Next, they were given detailed instructions for the DMS task. The instructions stressed the requirement to encode and maintain the stimuli in a holistic manner (as opposed to individual features) and to actively rehearse the stimuli across all three delay phases. Subsequently, participants were given 12 practice trials (three of each condition) to familiarize themselves with the task. Stimuli used in the practice trials were not used in the experiment proper. Next, participants received instructions for and performed one face block and one house block from the 1-back task (again using stimuli not presented in the scanner). Then, participants were taken to the MRI room and scanning commenced. All participants performed the DMS task first, followed by the 1-back task. After leaving the scanner, participants filled in a postexperimental questionnaire and were debriefed.

### Data Acquisition

Imaging was performed using a 3-T Siemens Magnetom Trio scanner (Siemens Medical Solutions, Erlangen, Germany). In both runs of the DMS task, 660 images with 28 axial slices (3.2 mm in-plane resolution, 3.3 mm slice thickness, 20% spacing) parallel to the AC-PC plane were acquired using a multiecho EPI sequence (Poser, Versluis, Hoogduin, & Norris, 2006) and a 32-channel head coil. The repetition time was 2,200 msec, the flip angle was  $90^\circ$ , and the echo times were 9.4, 21.2, 33, 45, and 57 msec. The field of view was  $205 \times 205 \text{ mm}^2$ . The first four images were discarded from the analysis. The next 20 images were rest scans that were later used to calculate weighting images (see below). Identical scanning parameters were used for the 1-back task, but only 190 images were acquired per run. In a separate scanning session, a high-resolution 3-D T1-weighted data set with  $1 \times 1 \times 1.25 \text{ mm}^3$  resolution was acquired.

### Data Analysis

#### *Behavioral Data Analysis*

Behavioral and ROI data were analyzed with IBM SPSS 22 (IBM Corp., Armonk, NY) and IPython notebook (Pérez & Granger, 2007). Unless otherwise noted, error bars in line and bar plots represent 95% within-subject confidence intervals (CIs; Morey, 2008; Cousineau, 2005). The calculations of Cohen's *d* disregarded correlations,



that is, the data were treated as between-subject data (Dunlap, Cortina, Vaslow, & Burke, 1996).

### *Preprocessing of fMRI Data*

fMRI data processing was carried out using FSL (FMRIB's Software Library; Smith et al., 2004). As we used a multi-echo sequence with five echoes, the reconstructed data consisted of five complete time series for each run, one for every echo time. In the first preprocessing step, these time series were combined into a single time series per run. To this end, we used the time series from the first echo to compute motion correction parameters using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002). These parameters were then applied to all time series. Next, the five time series were split into two parts, corresponding to the rest periods acquired at the beginning of each run (the "weighting time series") and the task proper (the "task time series"). The data corresponding to the weighting time series were then used to compute weighting images as described previously (Poser et al., 2006). Briefly, the weight of a voxel in a particular weighting image depends on its mean signal strength and variability relative to this voxel's values in the other weighting time series. The weight will be high if a voxel's signal is strong and stable in a particular time series relative to the other time series. One weighting image was computed for each weighting time series. Next, the five task time series were multiplied by their corresponding weighting images and added up to create a single time series that was used for the remainder of the analysis.

The following preprocessing steps were then applied within FEAT (version 5.98): nonbrain removal using BET (Smith, 2002), grand mean intensity normalization of the entire 4-D data set by a single multiplicative factor, and high-pass temporal filtering (Gaussian-weighted least squares straight line fitting, 128 sec). Finally, spatial smoothing was applied with a Gaussian kernel of 5 mm FWHM.

### *Modeling and Multivariate Analysis*

Following the approach proposed by Mumford, Turner, Ashby, and Poldrack (2012), the sample phase and all three delay phases of every single trial were modeled in separate general linear models (GLM) using a double-gamma function; for each such GLM, all other events contributed to a single regressor of no interest. The sample (i.e., memory encoding) phase was modeled by an epoch of 4.4 sec length. Unfilled delays (i.e., the pre-distractor delay and the post-distractor delay) were modeled by a 4.4-sec epoch placed in the middle of the 8.8-sec delay period (Zarahn, Aguirre, & D'Esposito, 1997). Distractor delays were modeled by 6.6-sec epochs. Motion correction parameters obtained during echo combination were used as confound regressors. For each participant, the parameter estimates (i.e., one beta image per trial period)

from Run 2 were then spatially registered to Run 1, based on transformation matrices obtained from registering the middle volume of Run 2 to the middle volume of Run 1 using six degrees of freedom and normalized correlation as a cost function. To prepare the data for the multivariate analysis, the parameter estimates from all trial phases of all correctly answered trials were concatenated in the order in which they were acquired. If a participant made no errors, this resulted in a 4-D file with 288 volumes (i.e., 72 trials, parameter estimates for sample phase and three delay phases). For each image in the concatenated file, an attribute file labeled its sample category (i.e., face or house) and its phase in the trial (i.e., sample, pre-distractor, distractor, post-distractor). A separate set of attribute files was created that also labeled the type of distraction (object or scrambled).

PyMVPA (Hanke et al., 2009; [www.py\\_mvpa.org/](http://www.py_mvpa.org/)) was then used for linear detrending and z-transforming the data for each cross-validation fold. A searchlight analysis (Kriegeskorte, Goebel, & Bandettini, 2006) with a radius of two voxels (33 voxels per searchlight) and a support vector machine (SVM) classifier with PyMVPA's default C parameter (which automatically scales C according to the norm of the data) with cross-validation was performed to identify voxels that distinguished between maintenance of face versus house information in the different delay phases. The chunks used for cross-validation were the blocks into which the runs were subdivided. Thus, for each participant, there were six chunks overall. For each cross-validation fold, five chunks were used for classifier training and the remaining chunk was used for classifier testing. For each fold of the cross-validation procedure, we made sure that an equal number of training samples was available for face and house trials (a different approach was used for the analysis that correlated behavioral and decoding performance—see below). An example might help to illustrate this: There was a maximum of 30 correct trials per sample category in five chunks (i.e., 5 blocks  $\times$  6 trials per condition and block). If a participant had made no house trial errors but two errors on face trials, only 28 face trials remained for training. The training procedure ensured that an equal number of house trials (i.e., 28) was then randomly chosen from the 30 available ones. To mitigate the effects of randomly selecting a subset for one category, we repeated this selection process five times for each fold and averaged the classifier performance across the five repetitions (given that differences in trial numbers were typically very small, further increasing the number of repetitions had very little effect on the results).

Classifier performance was evaluated by calculating the balanced accuracy averaged across the six folds. The balanced accuracy results from calculating accuracies per target category initially, before averaging over target categories. The advantage of this measure is that it is independent of the relative frequency of the target categories in a chunk used for testing. For example, if the classifier









**Table 3.** (continued)

Location	Volume (mm <sup>3</sup> )	z-Value	x,y,z
Lingual gyrus		4.97	12, -82, -12
Inferior temporal gyrus		4.93	56, -60, -18
Occipitotemporal cortex		4.81	58, -64, 18
Fusiform gyrus		4.73	30, -42, -20
Superior parietal lobule	8,312	3.76	-8, -70, 58
Superior parietal lobule		4.89	18, -60, 60
Cingulate sulcus		3.81	14, -36, 42
Parahippocampal gyrus	3,856	4.13	26, -6, -32
Temporal pole	3,840	5.15	-50, 22, -18
Rostral gyrus	2,424	3.74	-4, 46, -14

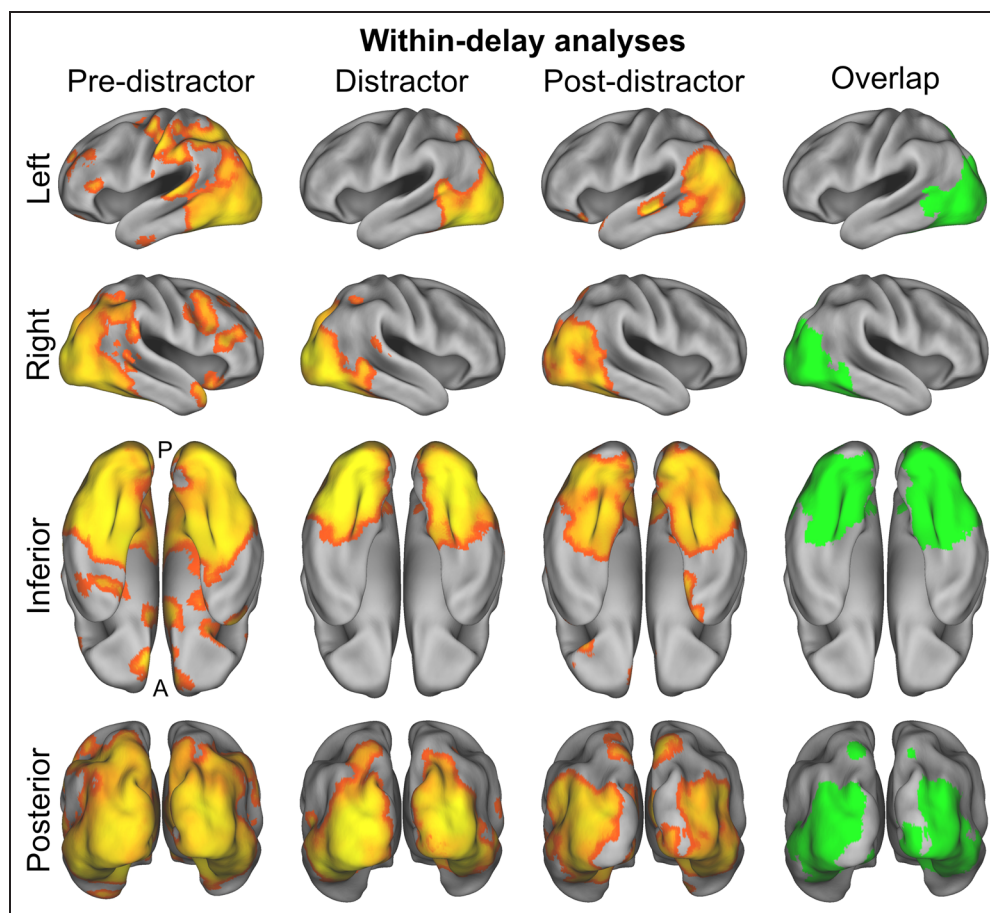
The  $z$  maps were thresholded using clusters determined by  $z > 3.1$  and a corrected cluster significance threshold of  $p < .05$ . Peaks are at least 25 mm apart. Up to 10 peaks per cluster are reported.

analyses would provide evidence that the neuronal code representing sample category information was either robustly maintained across distraction or reinstated after the distractor delay. We found that sample categories could be successfully decoded when training on the pre-distractor delay and testing on the post-distractor de-

lay, and vice versa. Regions with above-chance decoding in both analyses were found in the parahippocampal and fusiform gyri (Table 4; Figure 3A) and overlapped with functionally defined FFA and PPA (see below).

The areas identified in the present across-delay analyses overlapped with the anterior inferotemporal regions

**Figure 2.** Results of the within-delay searchlight analyses decoding STM representations of faces versus houses. For these analyses, a multivariate classifier was trained and tested on independent chunks of the same delay phase. Depicted in color are areas where decoding performance was above chance ( $z > 3.1$ , cluster threshold  $p < .05$ ). The rightmost column depicts the overlap in all three delay phases.



**Table 4.** Across-delay Analyses for the Pre- and Post-distractor Delays

Location	Volume (mm <sup>3</sup> )	z-Value	x,y,z
<i>Train on Pre-, Test on Post-distractor Delay</i>			
Collateral sulcus	18,400	4.91	-30,-38,-16
Fusiform gyrus		4.15	-30,-58,-14
Collateral sulcus	9,568	4.31	26,-38,-16
Parahippocampal gyrus		3.9	36,-22,-24
Lateral occipitoparietal cortex	2,096	3.68	-40,-84,26
<i>Train on Post-, Test on Pre-distractor Delay</i>			
Occipitotemporal sulcus	19,808	4.46	-44,-44,-10
Fusiform gyrus		4.29	-28,-44,-16
Inferior temporal sulcus		3.55	-54,-30,-10
Collateral sulcus	6,560	3.72	22,-42,-14
Fusiform gyrus		3.57	36,-24,-26

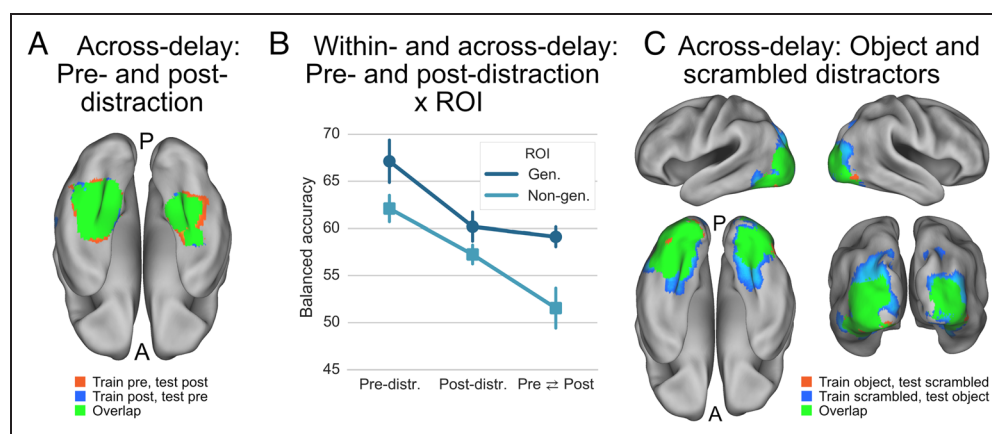
The  $z$  maps were thresholded using clusters determined by  $z > 3.1$  and a corrected cluster significance threshold of  $p < .05$ . Peaks are at least 15 mm apart.

identified in all three within-delay analyses. This suggests that only in ventral temporal areas the category-specific pattern of activation could survive or be reinstated, whereas the pattern of activation in more posterior areas was modified after the presentation of the distractors. To formally test this observation, we created two ROIs based on the overlap map shown in Figure 2, that is, an ROI consisting of those voxels showing significant across-delay generalization and another ROI consisting of all re-

maining voxels in the overlap map (i.e., that did not generalize from pre- to post-distractor delay). We performed a  $2 \times 3$  repeated-measures ANOVA on the mean balanced accuracies (cf. Methods) retrieved from these ROIs with factors ROI (generalization or nongeneralization) and Analysis (pre-distractor, post-distractor, or across-delay analysis).

The results showed a main effect of ROI,  $F(1, 20) = 46.8$ ,  $p < .001$ ,  $\eta_p^2 = .7$ , and a main effect of Analysis,

**Figure 3.** (A) Results of the across-delay analysis of face versus house representations involving the pre- and post-distractor delays. Depicted in color are areas where decoding performance was above chance ( $z > 3.1$ , cluster threshold  $p < .05$ ), indicating that in these areas the pattern of pre-distractor maintenance-related activity was reinstated after distraction. A ventral view of the PALS-B12 atlas brain is shown. (B) Results of an ROI analysis further investigating the reinstatement of maintenance-related activity patterns. The analyses labeled “Pre-distr.” and “Post-distr.” are the respective within-delay analyses. The analysis labeled “Pre  $\rightleftharpoons$  Post” is the across-delay analysis involving training and testing on the pre-distractor delay and the post-distractor delay. The ROIs were defined based on the overlap map shown in Figure 2. The generalization ROI (“Gen.”) corresponds to areas within the overlap map where evidence was reinstated after the distractor delay, the nongeneralization ROI (“Non-gen.”) corresponds to areas within the overlap map where evidence was not reinstated. The main result of this analysis is that the balanced classifier accuracy in the nongeneralization ROI drops significantly in the Pre  $\rightleftharpoons$  Post analysis relative to the post-distractor analysis, whereas this is not the case for the generalization ROI. Error bars represent 95% within-subject CIs. (C) Results of the across-delay analysis involving object and scrambled distractor delays. Depicted in color are areas where decoding performance was above chance ( $z > 3.1$ , cluster threshold  $p < .05$ ), indicating that these areas represented sample information in the distractor delay in a similar way irrespective of the type of distractor.



$F(2, 40) = 30.2, p < .001, \eta_p^2 = .6$  (Figure 3B). Crucially, there was also a significant interaction effect,  $F(2, 40) = 9.4, p < .001, \eta_p^2 = .59$ . For the across-delay analyses, the less accurate of the two within-delay analyses likely provides an upper limit for the generalization classifier performance. Therefore, in our post hoc analyses of the interaction effect, the across-delay results will be evaluated relative to the less accurate within-delay analysis, that is, the post-distractor delay. Note that these analyses are independent of how the ROIs were selected. In the non-generalization ROI, decoding performance was indeed worse for the across-delay analysis, compared with the post-distractor delay analysis,  $t(20) = 4.9, p < .001$ , Cohen's  $d = 1.17$ , whereas there was no significant difference in the generalization ROI,  $t(20) = 1.2, p = .23$ , Cohen's  $d = 0.15$ . As a  $t$  test cannot provide evidence for the null hypothesis, we also calculated the scaled JZS Bayes factor (Cauchy prior width = 0.707). The Bayes factor was 2.3, indicating that, given the data, the null hypothesis of no difference between the conditions resulted in being twice as likely, which can be considered weak evidence in favor of the null hypothesis. For the nongeneralization ROI, a control analysis showed that the drop in accuracy in the pre–post analysis (5.7%) was almost identical (5.6%) when voxels with lower accuracies in the post-distractor analysis were omitted from the analysis (and, as a result, the mean accuracy of the generalization and non-generalization ROIs in the post-distractor analysis was matched). This suggests that the observed pre–post analysis drop in accuracy in the nongeneralization ROI is not driven by a subset of voxels with lower overall accuracies or higher levels of noise. Taken together, these analyses indicate that in bilateral ventral temporal cortex (fusiform and parahippocampal gyri) the patterns of activity for the two categories were stable across the pre-distractor delay and the post-distractor delay (and thus a classifier trained on distinguishing these categories in one type of delay could also distinguish them in another), whereas in more posterior areas (involving occipital, posterior temporal, and parietal cortex) the patterns of activity changed over time, suggesting that in these latter areas the neural representations of memory contents were unique to the respective delay.

#### *Across-delay Analyses Involving the Distractor Delay*

As a third step, we tested whether or not generalization was still possible when the analyses involved the distractor delay. First, we trained on the pre-distractor delay and attempted to classify in the distractor delay, and vice versa. The searchlight analysis did not show any areas with above-chance decoding. Next, we trained on the distractor delay and attempted to classify in the post-distractor delay, and vice versa. Again, sample category could not be successfully decoded. An inspection of the mean balanced accuracies for the above analyses in the overlap areas shown in Figure 2 revealed that these accuracies

were slightly below chance (pre-distractor and distractor:  $M = 47.3\%, SD = 2.2$ ; post-distractor and distractor:  $M = 49.2\%, SD = 2.4$ ). To further explore the null effects in the decoding analyses, directional Bayes factors (alternative hypothesis:  $M > 50\%$ ; Cauchy prior width = 0.707) for the mean balanced accuracies in the overlap ROI were computed. The Bayes factors were 19.4 and 10, respectively, constituting strong evidence in favor of the null hypothesis. Thus, although it was possible to decode STM contents when training and testing were based on the distractor delay, decoding did not generalize between unfilled delay periods and the distractor delay period, suggesting that some information about the sample category is maintained during the distractor delay, but that the pattern of activity representing this information is fundamentally different from the other delays. This indicates that activation patterns representing STM contents in the ventral temporal cortex are not maintained across all delays, but are reinstated after distraction.

A potential problem of the previous across-delay analyses involving the distractor delay is that the distractor delay is the only delay in which stimuli were presented. Thus, for example, training on the pre-distractor delay and testing on the distractor delay involves training in the absence of perceptual input and testing in the presence of perceptual input. To address this issue, we repeated the analyses after training on the encoding phase. Using this approach, the classifier generalized to the pre- and post-distractor delay phases (in both directions). The regions found in these analyses overlapped with those identified in the above-reported within-delay analyses (results not shown). However, decoding for the distractor phase was still not successful. These results strengthen our conclusion that the pattern of activation representing information about samples is qualitatively different in the distractor delay.

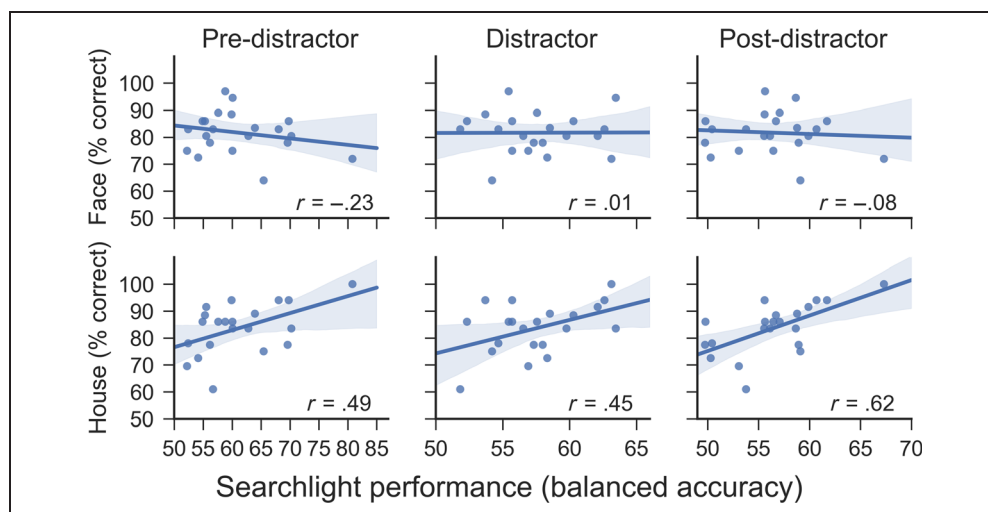
#### *Decoding Sample Category during Object and Scrambled Distraction*

To maximize power, our initial analysis collapsed across distractor type (i.e., object and scrambled distractors). As a result, successful decoding during the distractor phase might be dominated by one of the distractor types. A more rigorous demonstration of maintenance during the distractor delay would be to show that it is possible to train the classifier to distinguish the sample categories during one type of distraction and then decode them during the other. As Figure 3C shows (see also Table 5), this analysis was indeed successful. The regions common to both analyses were highly similar to the distractor-type independent analysis (Figure 2, second column) and involved the occipital cortex and the posterior fusiform gyrus. This suggests that, in these regions, a common pattern of activity is present during both types of distractor delay.





**Figure 4.** Analyses correlating classifier performance in the three different delay intervals with behavioral performance in face and house memory trials. Notably, behavioral accuracy in face trials is never correlated with classification accuracy, whereas behavioral accuracy in house trials is correlated with classification accuracy in all delay phases.



(as noted in the Methods section, the left FFA could only be identified when no cluster size correction was used) and restricted to the mid-fusiform sulcus (Weiner et al., 2014) and the right STS. House-preferential regions on the other hand were extensive, covering parts of the

parahippocampal gyrus (corresponding to the PPA), the medial fusiform gyrus, the inferior temporal gyrus, the occipital gyri, and the superior parietal lobule.

As apparent when comparing Figures 5A and 2, both face- and house-preferential areas overlapped with regions

**Table 6.** Results of the Functional Localizer Analysis

Location	Volume ( $\text{mm}^3$ )	z-Value	x,y,z
<i>Face vs. House</i>			
Mid-fusiform sulcus	240	4.3	-44,-46,-22
Mid-fusiform sulcus	1,400	4.61	44,-48,-18
Superior temporal sulcus	544	3.95	46,-46,18
<i>House vs. Face</i>			
Collateral sulcus	53,232	6.69	-26,-46,-8
Lateral occipital gyrus		6.42	-38,-88,14
Fusiform gyrus		6.39	-26,-80,-14
Paroccipital segment of the intraparietal sulcus		5.74	-26,-76,40
Parieto-occipital fissure		5.02	-16,-62,12
Inferior temporal gyrus		4.99	-50,-62,-6
Occipital pole		4.71	-18,-100,0
Collateral sulcus	53,280	7.42	24,-46,-10
Lateral occipital gyrus		6.61	34,-84,20
Superior parietal lobule		6.49	20,-72,52
Parieto-occipital fissure		6.31	18,-58,16
Fusiform gyrus		5.45	26,-88,-10
Inferior temporal gyrus	632	4.3	52,-60,-12

The z maps were thresholded using clusters determined by  $z > 3.1$ . A corrected cluster significance threshold of  $p < .05$  was applied to the house versus face map. No cluster correction was applied to the face versus house map (see Methods). In addition, both maps were exclusively masked by contrasts against the implicit baseline. Peaks are at least 25 mm apart.



and decoding accuracy in the distractor delay were significantly correlated,  $r_{Y(1,2)}(21) = .62, p = .004$ . This suggests that, whereas suppression in house trials has an influence on the decoding accuracy, the suppression effect does not drive the correlation of behavioral and decoding accuracy.

## DISCUSSION

This study investigated the neural correlates of distractor-resistant STM employing a DMS task with face or house stimuli. The task had three delay phases, a pre-distractor, a distractor, and a post-distractor delay. Using fMRI and multivariate searchlight analyses, we found that a support vector machine classifier could successfully distinguish face from house maintenance in all three delay phases when trained and tested on the same delay phase (within-delay analyses). Thus, our results provide support for active maintenance in all three delay intervals. Using across-delay analyses (where the classifier was trained on one delay and tested on another), we found that the classifier generalized from the pre-distractor delay to the post-distractor delay (and vice versa) in ventral temporal lobe areas overlapping with functionally defined face- and house-preferential areas. However, generalization failed when the training or the testing data set included the distractor delay, suggesting that patterns of maintenance-related activity were different when distractors were present. Our task employed two different types of distractors, scrambled and object pictures. Results showed that it was possible to train on one type of distractor and decode the memory content while the other distractor type was presented. This result suggests that the type of distractor had no major effect on how memoranda were maintained in the distractor delay.

In a correlational analysis, we were furthermore able to demonstrate that behavioral performance in house memory trials was positively correlated with decoding accuracy in posterior cortical areas. This suggests that, for house trials, the patterns of activity in the different delay intervals reflected how well sample-related information was maintained. Presumably, these patterns allowed the classifier to distinguish between the sample categories, leading to improved classifier performance, and allowed the participant to more precisely match sample and probe, leading to improved behavioral performance. Another correlational analysis showed that the success of classifier generalization in ventral temporal areas where activity patterns were recovered after distraction was also correlated with performance in house trials. This suggests that for these temporal areas the fidelity of recovery was relevant for task performance. It is noteworthy that no analogous brain-behavior correlation was observed for performance on face maintenance trials. We discuss possible reasons for this differential finding below (see Limitations and Open Questions section).

The remainder of our discussion will initially focus on how our study relates to previous STM research (focusing on distractor resistance) and what it contributes to this research. We will then move on to a discussion of possible reasons as to why some previous studies failed to find evidence for active maintenance in distractor delays. Finally, we will discuss limitations of our approach and open questions.

### Relation to Previous Studies

Miller et al.'s (1996) classic study showed that lateral prefrontal neurons in macaques are involved in the recovery of STM contents after distraction. However, this study did not address the question if STM contents are actively maintained during distraction. Recently, Jacob and Nieder (2014) showed that parietal area VIP neurons do maintain such distractor-resistant representations in a numerosity task, thus highlighting the role of posterior cortices in distractor-resistant memory. Our fMRI results obtained in humans concur with these results in that they also implicate posterior cortices. However, using a task that required maintenance of information about visual appearance (as opposed to numerosity), we found that more posterior and ventral areas were involved in STM maintenance during distraction.

This study differs from previous fMRI studies investigating distractor-resistant STM (Bettencourt & Xu, 2016; Jha et al., 2004; Sakai, Rowe, & Passingham, 2002b) by having three extended delay phases, which allowed us to use MVPA not only to study delay-specific STM maintenance but also to study the generalization of activity patterns across delays. The results showed that the presence of distractors altered the neural code for sample maintenance in a way that classifiers that could decode STM contents during maintenance-only (and, in a control analysis, during stimulus encoding) could not decode the contents of STM in the presence of distractors. This suggests that the pattern of maintenance-related activity is modified when STM contents have to be shielded against distraction. Generally, STM is thought to operate by maintaining activity that is similar to the activity originally elicited by the sample (D'Esposito & Postle, 2015). Thus, the presentation of distractors might elicit activity similar to the maintained samples, and this might lead to interference between the neural representations of the information in memory and the distractor. One strategy to deal with this may be to alter the maintenance-related neuronal code in a way that makes it less similar to activity elicited by distractors.

Another important difference between our study and previous studies investigating distractor-resistant maintenance is that we required participants to actively process the distractor pictures, whereas the studies by Bettencourt and Xu (2016) and by Jha et al. (2004) used passive distractor conditions where participants were instructed to simply watch the distractors. We chose an active distractor





which we attribute to the following reasons: (1) The samples on each trial were unique and had never been seen by the participants before—thus precluding familiarity-based task performance. (2) Distractors needed to be attended and were presented in the same location as the samples. Therefore, lower-level stimulus representations were presumably “overwritten,” which might have impeded passive STM based on lower-level perceptual representations. (3) About 25% of the probe picture were covered by a black bar at an unpredictable location and nonmatch probe pictures were selected from the set of pictures previously presented as samples. Again, these factors presumably made the use of a passive strategy less likely to be successful. (4) Memory load was relatively low (2 items), being well within the typical working memory capacity ( $4 \pm 1$  items; Cowan, 2001).

Some of these factors might explain why, as mentioned above, a recent study by Lewis-Peacock et al. (2012) was not successful at decoding memory content. In their Experiment 1, distractors always came from a category that was not currently maintained. Moreover, non-match probes were not different exemplars of the same category but came from an irrelevant category. Furthermore, there were only 18 stimuli overall which had been learned before the fMRI scan, potentially allowing the formation of long-term memory representations. All of these factors may have reduced the likelihood that participants relied on active maintenance. In a second experiment by the same authors (Experiment 2), two samples were spatially separated, and apart from cues, no further stimuli were presented in the location of the samples. This might have left lower-level perceptual representations mostly intact and thus have facilitated reliance on passive memory mechanisms.

None of the factors discussed as possible reasons for the inability to decode STM contents in Lewis-Peacock et al. (2012) apply to the study by Sakai et al. (2002b), raising the question why these authors observed no activation differences between correct and incorrect trials in the distractor delay. One possible explanation is that the authors applied a purely univariate approach to data analysis, which might be less sensitive than current multivariate approaches. However, there is also a potentially more interesting explanation based on the nature of their tasks: In Sakai et al.’s study, the combined memory load from main task and distractor task was 10 items, which is considerably above the typical working memory capacity (Cowan, 2001) and makes it very unlikely that an active memory strategy could have succeeded.

### Limitations and Open Questions

Some of our results suggest that faces and houses were treated differently by our participants. Performance in face and house trials was only weakly correlated, and decoding accuracy was only related to behavioral house, but not face trial performance. In a postexperiment question-

naire, 59% of participants reported that one of the strategies they employed was to try to associate sample faces with persons known to them (e.g., friends or celebrities); an analogous strategy was only reported by 14% of participants in the house condition. As a result, although all of the samples were new to participants, long-term memory representations might have played a more important role for face trials. This might explain why decoding accuracy was correlated with behavioral accuracy for house trials, but not for face trials.

Our results suggest that posterior cortices play a role in the active maintenance of information in long distractor delays. Currently, few functional imaging studies have investigated this issue and a number of open questions remain. For example, it is unclear for how long active maintenance during distraction is possible. In our task, the distractor delay was 6.6 sec (using an active distractor task), and in the study by Bettencourt and Xu (2016), it was 10.2 sec (using passive distraction). Furthermore, it is uncertain which factors determine if active maintenance is used or not. The timing of the task, the memory load, the material to be remembered, the preexistence of long-term memory representations, and the participant strategy might influence the results. Future research should also continue to investigate which areas contribute to distractor-resistant STM. Although Bettencourt and Xu (2016) stressed the role of the posterior intraparietal sulcus, we could successfully decode in occipital, temporal, and parietal cortices.

### Conclusion

This study investigated the maintenance of visual information in STM in unfilled and distractor-filled delays. A multivariate searchlight analysis successfully decoded STM contents (faces or houses) in all delay phases. Regions with above-chance decoding in all delay phases were located in the occipital, temporal, and posterior parietal lobes. In ventral temporal cortex, including functionally defined areas FFA and PPA, activity patterns were reactivated after distraction. In more posterior regions, activity patterns were more flexible and depended on the delay phase. Classifier performance in all delay phases was correlated with the behavioral performance in house trials, but not face trials. The present results highlight the role posterior cortices play in the online maintenance of STM contents both in the presence and absence of distractors.

### Acknowledgments

This research was supported by a VIDI grant from the Netherlands Organization for Scientific Research (Grant 452-09-006 to C. J. F.). J. D. was further supported by the German Research Foundation (Grant DE 1141/2-1). M. H. was supported by funds from the German Federal State of Saxony-Anhalt and the European Regional Development Fund (ERDF; project: Center for Behavioral Brain Sciences). M. T. was supported by the German Research

Foundation in the Transregional Collaborative Research Center 134. C. J. F. was further supported by an ERC consolidator grant (No. 617891).

Reprint requests should be sent to Jan Derrfuss, School of Psychology, University of Nottingham, Nottingham, NG7 2RD, UK, or via e-mail: jan.derrfuss@nottingham.ac.uk.

## REFERENCES

- Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A., & Gee, J. C. (2011). An open source multivariate framework for n-tissue segmentation with evaluation on public data. *Neuroinformatics*, *9*, 381–400.
- Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience*, *19*, 150–157.
- Björnsdotter, M., Rylander, K., & Wessberg, J. (2011). A Monte Carlo method for locally multivariate brain mapping. *Neuroimage*, *56*, 508–516.
- Bloemendaal, M., van Schouwenburg, M. R., Miyakawa, A., Aarts, E., D'Esposito, M., & Cools, R. (2015). Dopaminergic modulation of distracter-resistance and prefrontal delay period signal. *Psychopharmacology*, *232*, 1061–1070.
- Clapp, W. C., Rubens, M. T., & Gazzaley, A. (2010). Mechanisms of working memory disruption by external interference. *Cerebral Cortex*, *20*, 859–872.
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*, 42–45.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114.
- D'Esposito, M., Cooney, J. W., Gazzaley, A., Gibbs, S. E. B., & Postle, B. R. (2006). Is the prefrontal cortex necessary for delay task performance? Evidence from lesion and fMRI data. *Journal of the International Neuropsychological Society*, *12*, 248–260.
- D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142.
- Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*, 170–177.
- Endl, W., Walla, P., Lindinger, G., Lalouschek, W., Barth, F. G., Deecke, L., et al. (1998). Early cortical activation indicates preparation for retrieval of memory for faces: An event-related potential study. *Neuroscience Letters*, *240*, 58–60.
- Ester, E. F., Rademaker, R. L., & Sprague, T. C. (2016). How do visual and parietal cortex contribute to visual short-term memory? *eNeuro*, *3*, e1.
- Etzel, J. A., Zacks, J. M., & Braver, T. S. (2013). Searchlight analysis: Promise, pitfalls, and potential. *Neuroimage*, *78*, 261–269.
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neuroscience*, *8*, 1298–1300.
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, *7*, 37–53.
- Jacob, S. N., & Nieder, A. (2014). Complementary roles for primate frontal and parietal cortex in guarding working memory from distractor stimuli. *Neuron*, *83*, 226–237.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*, 825–841.
- Jha, A. P., Fabian, S. A., & Aguirre, G. K. (2004). The role of prefrontal cortex in resolving distractor interference. *Cognitive, Affective & Behavioral Neuroscience*, *4*, 517–527.
- Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences, U.S.A.*, *103*, 3863–3868.
- Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*, 61–79.
- Lewis-Peacock, J. A., & Postle, B. R. (2012). Decoding the internal focus of attention. *Neuropsychologia*, *50*, 470–478.
- Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, *16*, 5154–5167.
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64.
- Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage*, *59*, 2636–2643.
- Oldfield, R. C. (1971). Assessment and analysis of handedness—Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Pérez, F., & Granger, B. E. (2007). IPython: A system for interactive scientific computing. *Computing in Science & Engineering*, *9*, 21–29.
- Piekema, C., Kessels, R. P., Rijpkema, M., & Fernandez, G. (2009). The hippocampus supports encoding of between-domain associations within working memory. *Learning and Memory*, *16*, 231–234.
- Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine*, *55*, 1227–1235.
- Race, E., LaRocque, K. F., Keane, M. M., & Verfaellie, M. (2013). Medial temporal lobe contributions to short-term memory for faces. *Journal of Experimental Psychology-General*, *142*, 1309–1322.
- Reddy, L., & Kanwisher, N. (2006). Coding of visual objects in the ventral stream. *Current Opinion in Neurobiology*, *16*, 408–414.
- Sakai, K., & Passingham, R. E. (2004). Prefrontal selection and medial temporal lobe reactivation in retrieval of short-term verbal information. *Cerebral Cortex*, *14*, 914–921.
- Sakai, K., Rowe, J. B., & Passingham, R. E. (2002a). Parahippocampal reactivation signal at retrieval after interruption of rehearsal. *Journal of Neuroscience*, *22*, 6315–6320.
- Sakai, K., Rowe, J. B., & Passingham, R. E. (2002b). Active maintenance in prefrontal area 46 creates distractor-resistant memory. *Nature Neuroscience*, *5*, 479–484.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*, 143–155.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*(Suppl. 1), S208–S219.
- Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring latent visual working memory representations in human cortex. *Neuron*, *91*, 694–707.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*, 751–754.
- Van Essen, D. C. (2005). A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage*, *28*, 635–662.

- Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, *8*, 443–459.
- Warrington, E. K., & Taylor, A. M. (1973). Immediate memory for faces: Long- or short-term memory? *Quarterly Journal of Experimental Psychology*, *25*, 316–322.
- Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., et al. (2014). The mid-fusiform sulcus: A landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage*, *84*, 453–465.
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage*, *14*, 1370–1386.
- Xu, Y., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*, 91–95.
- Yoder, P. J., Blackford, J. U., Waller, N. G., & Kim, G. (2004). Enhancing power while controlling family-wise error: An illustration of the issues using electrocortical studies. *Journal of Clinical and Experimental Neuropsychology*, *26*, 320–331.
- Yoon, J. H., Curtis, C. E., & D'Esposito, M. (2006). Differential effects of distraction during working memory on delay-period activity in the prefrontal cortex and the visual association cortex. *Neuroimage*, *29*, 1117–1126.
- Zanto, T. P., & Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. *Journal of Neuroscience*, *29*, 3059–3066.
- Zarahn, E., Aguirre, G., & D'Esposito, M. (1997). A trial-based experimental design for fMRI. *Neuroimage*, *6*, 122–138.
- Zlatkina, V., & Petrides, M. (2014). Morphological patterns of the intraparietal sulcus and the anterior intermediate parietal sulcus of Jensen in the human brain. *Proceedings of the Royal Society, Series B, Biological Sciences*, *281*, e1.