Journal Pre-proof

Content or status: Frontal and posterior cortical representations of object category and upcoming task goals in working memory

Katya Olmos-Solis, Anouk M. van Loon, Christian N.L. Olivers

Please cite this article as: Olmos-Solis K, van Loon AM, Olivers CNL, Content or status: Frontal and posterior cortical representations of object category and upcoming task goals in working memory, *CORTEX*, https://doi.org/10.1016/j.cortex.2020.11.011.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Katya Olmos-Solis: conceptualization, investigation, software, formal analysis, writing – original draft, writing - review & editing

Anouk M. van Loon: conceptualization, methodology, supervision, software, formal analysis, writing - review & editing

Christian N. L. Olivers: conceptualization, methodology, writing - review & editing, supervision, funding acquisition

Content or status:

Frontal and posterior cortical representations of object category and upcoming task goals in working memory

Katya Olmos-Solis[1],

Anouk M. van Loon[2]

&

Christian N. L. Olivers[1]

[1]*Institute for Brain and Behaviour, Vrije Universiteit Amsterdam, The Netherlands*

[2] *Department of General Practice & Elderly Care Medicine, Amsterdam Public Health Research Institute, Amsterdam University Medical Center, location VUmc, Amsterdam, The Netherlands*

Please address correspondence to

Prof C.N.L. Olivers

Institute for Brain and Behaviour Amsterdam

Section Cognitive Psychology

Faculty of Behavioural and Movement Sciences

Vrije Universiteit Amsterdam

The Netherlands

Email: c.n.l.olivers@vu.nl

Content or status:

Frontal and posterior cortical representations of object category and upcoming task goals in working memory

Katya Olmos-Solis[1],

Anouk M. van Loon[2]

&

Christian N. L. Olivers[1]

[1]*Institute for Brain and Behaviour, Vrije Universiteit Amsterdam, The Netherlands*

[2] *Department of General Practice & Elderly Care Medicine, Amsterdam Public Health Research Institute, Amsterdam University Medical Center, location VUmc, Amsterdam, The Netherlands*

Please address correspondence to

Prof C.N.L. Olivers

Institute for Brain and Behaviour Amsterdam

Section Cognitive Psychology

Faculty of Behavioural and Movement Sciences

Vrije Universiteit Amsterdam

The Netherlands

Email: c.n.l.olivers@vu.nl

Abstract

To optimize task sequences, the brain must differentiate between current and prospective goals. We previously showed that currently and prospectively relevant object representations in working memory can be dissociated within object-selective cortex. Based on other recent studies indicating that a range of brain areas may be involved in distinguishing between currently relevant and prospectively relevant information in working memory, here we conducted multivoxel pattern analyses of fMRI activity in additional posterior areas (specifically early visual cortex and the intraparietal sulcus) as well as frontal areas (specifically the frontal eye fields and lateral prefrontal cortex). We assessed whether these areas represent the memory content, the current versus prospective status of the memory, or both. On each trial, participants memorized an object drawn from three different categories. The object was the target for either a first task (currently relevant), a second task (prospectively relevant), or for neither task (irrelevant). The results revealed a division of labor across brain regions: While posterior areas preferentially coded for content (i.e. the category), frontal areas carried information about the current versus prospective relevance status of the memory, irrespective of the category. Intraparietal sulcus revealed both strong category- and status-sensitivity, consistent with its hub function of combining stimulus and priority signals. Furthermore, cross-decoding analyses revealed that while current and prospective representations were similar prior to search, they became dissimilar during search, in posterior as well as frontal areas. The findings provide further evidence for a dissociation between content and control networks in working memory.

Keywords: Multivoxel pattern analysis, working memory, visual search, multiple tasks, prospective memory.

Introduction

Humans often perform sequences of tasks for which they have to memorize multiple objects. A typical example is having a mental list of what to buy in the supermarket. Although all items in mind will become task-relevant at some point, in each step of the sequence only one item should guide our perceptions and actions, while all other items still need to be remembered, but without interfering with the current task. Visual working memory (VWM) is considered to be crucial for the maintenance of task-relevant representations in service of visual tasks and a vast body of evidence indicates that both sensory (Emrich, Riggall, LaRocque, & Postle, 2013; Ester, Serences, & Awh, 2009; Harrison & Tong, 2009; Riggall & Postle, 2012; Sreenivasan, Vytlacil, & D'Esposito, 2014) and fronto-parietal cortices (Christophel, Hebart, & Haynes, 2012; D'Esposito & Postle, 2015; Postle et al., 2006; Rainer, Asaad, & Miller, 1998; Ranganath, Cohen, Dam, & D'Esposito, 2004) participate in the process of either storing or manipulating relevant visual representations. However, much of the knowledge has been gained from single task settings. In order to optimize attentional deployment and ensure efficient behavior during task *sequences*, the brain must differentiate between information required for current versus prospective goals. Little is known about how the brain distinguishes between memories relevant for now and memories relevant for the near future.

Several hypotheses have been raised on how the brain may differentiate between currently relevant and prospectively relevant representations. One possibility is that, in contrast to memories that are currently relevant, prospectively relevant memories are stored 'silently', through changes in responsivity – rather than the activity – of the network that represents them (Olivers, Peters, Houtkamp, & Roelfsema, 2011; Stokes, 2015). These changes in responsivity may happen through short-term synaptic plasticity (Mongillo, Barak, & Tsodyks, 2008; Sugase-Miyamoto, Liu, Wiener, Optican, & Richmond, 2008) or changes in the membrane potentials of the previously activated neurons (Conde-Sousa & Aguiar, 2013; Stokes, 2015). Consistent with this hypothesis, prospectively relevant items are often not decodable even though participants are still able to remember them (Larocque, Lewis-Peacock, & Postle, 2014; Larocque, Riggall, Emrich, & Postle, 2017; Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012), but can be successfully recovered when unrelated activity is sent through the network, either as a burst of transcranial brain stimulation (TMS) (Rose et al., 2016) or by presenting other visual input unrelated to the memorized item (Wolff, Jochim, Akyürek, & Stokes, 2017). In addition, we recently found evidence that such silent prospective memories may be stored in a transformed, 'inverted' format relative to current memories (van Loon, Olmos-Solis, Fahrenfort, & Olivers, 2018; Yu, Teng, & Postle, 2020).

A different hypothesis comes from Christophel, Iamshchinina, Yan, Allefeld, and Haynes (2018), who have argued that prospective memories are not so much silently stored in sensory areas, but instead may resort to higher brain areas that are more robust against interference. They found that while currently relevant items could be decoded from primary sensory cortices (i.e., V1-V3), evidence for prospective memories was at baseline levels in those visual areas. Instead, prospective memories could be successfully decoded from parietal and frontal cortices, specifically intraparietal sulcus (IPS) and frontal eye fields (FEF). This suggests that the brain distinguishes between current and prospective memories by recruiting different brain regions for their storage. The transfer of prospective representations from sensory cortices to the frontoparietal network may be a more general functional mechanism put in place to prevent sensory-memory interference. Indeed, Bettencourt and Xu (2016) reported that under conditions of distraction during the delay period between the memorandum and the test, VWM representations appeared to be momentarily withdrawn from occipital cortex to parietal cortices (see however Rademaker, Chunharas, & Serences (2018) and Lorenc, Sreenivasan, Nee, Vandenbroucke, & D'Esposito (2018) for contrasting results). A similar transfer may also occur for prospective memories during task sequences, given that they need to be shielded from the potential interference of the current task.

How can the evidence in favor of these different hypotheses be reconciled? One possibility is that multiple brain systems participate in the process of representing prospective memories, but they do so differently, depending on how susceptible to perceptual interference a given brain area is. Specifically, while sensory areas may store prospective memories in a silent, transformed (i.e. inverted) format to prevent interference from upcoming sensory input, task-related representations in the fronto-parietal network are likely to be more robust against such sensory interference. Fronto-parietal areas could thus store task-related memories in a more active format regardless of their current or prospective status. If so, an additional question then is whether current and prospective representations are similar, or are dissociated within this network, through respectively uncorrelated (Warden & Miller, 2007) or inverted (Van Loon et al., 2018; Yu, Teng, & Postle, 2020) coding.

Another possibility is that different brain areas, rather than all representing memory content, code for different aspects of the task at hand. While posterior (i.e., more sensory) regions may preferentially code for the identity and features of memory contents, frontal areas may be less involved in representing the content itself, but rather reflect the goal, function or associated action of the memory. In other words, such areas may reflect the status (i.e. current versus prospective) of the memory, rather than its content or identity (cf. Lee, Kravitz, & Baker, 2013). Certainly, VWM

4

contents in general and prospective representations in particular have been decoded from frontal regions; especially when the studies have a sufficiently large sample size and statistical power (Christophel et al. 2018). However, previous studies have also shown that less than half of the prefrontal neurons are stimulus selective and even those neurons that are selective show less than 50% modulation of their firing rate by information in VWM (Miller, Erickson, & Desimone, 1996; Parthasarathy et al., 2017). Neurons that are not stimulus-selective could then serve to code for more general task rules which need to dynamically change according to task demands. In addition, note that there is also the distinct possibility that any content-related decoding from frontal areas reflects motor-related components. Especially orientation-related activity in the FEF may reflect stimulus-specific eye-movements, rather than the content itself, as has been shown for example for orientation information (Mostert et al., 2018).

To provide additional evidence on how currently and prospectively relevant memories are represented across different brain regions, we analyzed data acquired during a previously reported fMRI experiment (van Loon et al., 2018, Experiment 2), but here focused on four additional regions of interest (ROIs). Figure 1 shows the task design. In this experiment, observers first committed to memory a visually presented object drawn from one of three possible categories. The object was the target for one of two consecutive visual search tasks which were presented after an initial memory delay. Importantly, a cue indicated whether the item of interest was relevant for the first search (*current* relevance), relevant for the second search (*prospective* relevance), or if it was not relevant for either search task (the *irrelevant* condition). Our previous analyses focused solely on the posterior fusiform part of lateral occipital cortex (pFs), as this regions is known to represent object categories. Multivoxel pattern analyses (MVPA) of this area suggested that the memorized items underwent a marked representational transformation throughout the trial. Using a cross-relevance decoding scheme (i.e., training the classifier in current trials and testing it on prospective trials and vice-versa), we could directly compare the neural representation of the same categories when they were the target for the first search (i.e. currently relevant) versus when they were the target for the second search (i.e. prospectively relevant). The results revealed that during the delay prior to search, the neural representation of a category when prospective was similar to the representation of the same category when current. However, during the first search task, current and prospective representations of the same category became very dissimilar, to the extent that the neural pattern of a category when prospective was partially opposite to its pattern when current. The standard activity pattern then re-emerged when the prospective item became relevant again, for the second search task. We interpreted this inversion of the category-related activity patterns as a potential

mechanism to deal with sensory interference (caused by the categories presented during the search display) and prevent the prospective representation from interacting with the task at hand.

Three main questions guided our analyses and choice of ROI in the present study. First, we were interested in whether brain regions other than pFs carry information about the content (i.e. category membership) of currently relevant and prospectively relevant objects in memory and whether such category representation differed depending on the status. Second, if a brain region codes for object category, does it do so in the same way for current and prospective items? Specifically, we investigated if other areas would also reveal the representational transformation of category representations that we previously observed in pFs (Van Loon et al., 2018). Third, we evaluated whether the different ROIs can differentiate between current, prospective and irrelevant items, *regardless* of their category. In other words, whether those areas code for the level of relevance of the memorized item. To this end, we performed MVPA on data from visual cortex (V1- V3 combined), IPS, and FEF, and lateral prefrontal cortex (lPFC) to investigate if current and prospective object categories are represented within these areas and to what extent they differ. These ROIs were chosen on the basis of previous work showing that they are involved in visual working memory tasks and show task-specific modulations. Specifically, based on findings by Bettencourt and Xu (2016) and Christophel et al (2018), we might expect prospective information to be withdrawn from visual cortex and stored in higher areas instead, specifically IPS and FEF. In contrast, Rademaker, Chunharas, & Serences (2018) and Lorenc, Sreenivasan, Nee, Vandenbroucke, & D'Esposito (2018) showed that visual cortex still contained information on the memorandum in the wake of interfering stimuli, and we might expect the same to occur for prospective information while observers prepared for, and conducted, an intermediate search task. Given the involvement of specifically prefrontal cortex in the planning of behavior (Mansouri, Koechlin, Rosa, & Buckley, 2017), as well as in task-dependent modulations of visual working memory content (Lee, Kravitz, and Baker, 2013), we hypothesized that lPFC would be sensitive to whether a memory served the current or the prospective behavioral goal.

Materials and methods

We show previously unreported analyses of a data set reported on in an earlier publication, which focused on object-selective cortex (van Loon et al. 2018; Experiment 2). Here we analyzed four additional ROIs: early Visual Cortex (VC), Intra Parietal Sulcus (IPS), Frontal Eye Fields (FEF) and lateral Prefrontal Cortex (lPFC). We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data

analysis, all manipulations, and all measures in the study. The site of the earlier publication https://elifesciences.org/articles/38677/figures and in addition Open Science Framework page https://osf.io/hcp47/ contain the raw data, the first level GLM output data files per ROI, preprocessing scripts, decoding analyses scripts, as well as stimulus presentation scripts. All ROIs were chosen on the basis of previous studies (see above, and below) and all preprocessing steps, choice of dependent measures, and analyses followed van Loon et al. (2018), with the exception of applying a more rigorous classifier balancing procedure here (see below). None of procedures and analyses were preregistered.
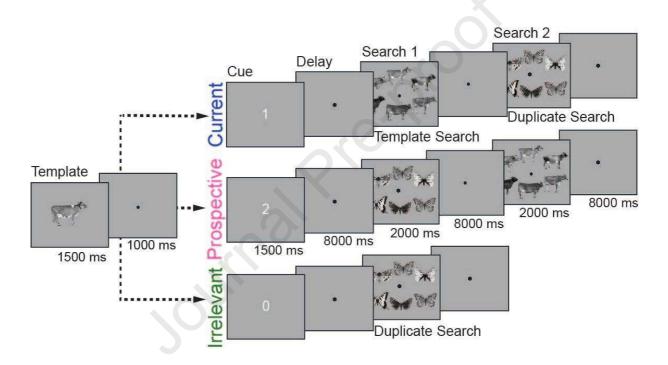

Figure 1. Trial design. On each trial, participants performed two consecutive visual search tasks. The target for one of those tasks was presented at the start of the trial and could be a cow, a dresser or a skate (variable template search; 4 exemplars per category). These categories were the basis for the category classification analyses. Then a retro-cue appeared, when the cue was "1" the memorized object was a current template, for Search 1; cue "2" indicated that the object was a prospective template, for Search 2; finally, when the cue was "0" the memorized item was not a target in either search and thus it was irrelevant in the trial. The remaining search task (either Search 2 in the Current condition, or Search 1 in the Prospective and Irrelevant conditions) was a so-called duplicate search task. In this task, butterflies, motorcycles or trees were presented and participants indicated whether or not any one of the exemplars was shown twice in the display. Thus, here no search template could or needed to be prepared. In the irrelevant condition participants only performed the duplicate search task. The main conditions (i.e., Current, Prospective or Irrelevant) formed the basis for the status decoding analysis. For each search display, participants indicated whether the target object was present

or absent using a button press. At the end of each run participants received feedback about their performance. The figure is the same as used in van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. L. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *eLife*, *7*, e38677.

Participants

Twenty-five observers participated in Experiment 2 of van Loon et al. (2018; 14 males, M = 25 years of age, SD = 4.5 years). Pre-determined exclusion criteria were a lack of normal or corrected-to-normal vision, plus the standard exclusion criteria for fMRI research (no metal, no pregnancy, etc.). We obtained written informed consent from each participant before experimentation. Sample size was determined on the basis of Experiment 1 in Van Loon et al., 2018), where highly reliable category and relevance decoding was found with N=24. We thus planned a sample size of 24. One subject did not complete the experiment and was therefore a priori excluded. An extra subject was tested in the remaining sessions to ensure that we would have a complete sample in case another participant had to be excluded. In the end, this turned out unnecessary, resulting in N=25. The experimental procedure was approved by the Ethical Committee of the Faculty of Social and Behavioral Sciences, University of Amsterdam (where scanning took place) and conformed to the Declaration of Helsinki.

Task and stimuli

The task design is shown in Figure 1. On each trial, participants performed two consecutive visual search tasks of real-world objects. The object of interest consisted of real-world greyscale photographs, selected out of four exemplars from three categories (cows, dressers or skates). These categories were selected to have maximal dissimilarity in representational space (see Harel, Kravitz, & Baker, 2014). The memorized object (cow, dresser or skate) was to be searched for first or second – thus making it currently or prospectively relevant (referred to as the variable template search). The remaining search task was a duplicate search where participants had to indicate if one of the objects appeared twice in the search display. This duplicate search tasks changed the category from trial to trial to be one out of three possible categories (butterfly, motorcycles and trees). Note that the duplicate search did not require a template, because all the information needed to perform the task was in the search display itself. As a result, observers only had to memorize a single item per trial, which was either the target for the first search (Current; with the second search being a duplicate search), the target for the second search (Prospective; with the first search being the duplicate search), or was deemed irrelevant after all (Irrelevant condition; with the first and only search then being a duplicate search).

Each trial started with the presentation the memory item (cow, dresser or skate) for 1500 ms (size: 2.4° visual angle), followed by a fixation display that stayed on for 1500 ms. Then, a cue indicated the relevance of this item. The cue could be either a 1, 2 or 0 and remained on the screen for 1000 ms. When the cue was '1', participants performed the template search first and the duplicate search second, making the memorized object currently relevant (Current condition). The order reversed when the cue was '2', rendering the memorized object only prospectively relevant, as observers performed the duplicate search first and the template search second (Prospective condition). Finally, if the cue was '0' the memorized object was irrelevant because participants would only perform the duplicate search (Irrelevant condition) and would not be tested on the object. The cue was followed by an 8 s delay interval with a fixation cross in the middle of the screen, after which the first search display was presented. Depending on the condition, the first search display was either a template search (Current condition) or a duplicate search (Prospective and Irrelevant conditions). In the template search, participants indicated with a button press whether the memorized object of interest was present or absent among six exemplars of the same object category. Similarly, in the duplicate search, participants indicated whether a duplicate object (i.e., the same exemplar appeared twice in the search display) was present or absent, again set size for this display was six objects. The objects in the search displays were randomly placed on the virtual rim of a circle with a radius of 7.4° of visual angle. They were presented for two seconds and observers had to respond within these two seconds. In all conditions, after the first search display, another eight seconds blank period followed. Then, the trial either ended (Irrelevant condition) or the second search display was presented (Current and Prospective conditions). This second search display was also followed by an eight second blank period after which the trial ended.

The experiment consisted of 9 runs with 12 trials each (108 trials in total). Within each experimental run, we balanced the amount of times that each relevance condition (Current, Prospective and irrelevant) was presented (four trials per condition), as well as the amount of times that participants had to respond either 'present' or 'absent' in each search task. However, the relevance condition by category combinations (i.e., nine in total: three relevance conditions [current, prospective, irrelevant] x three memory category [cow, dresser, skate]) could not be completely balanced within runs (12 trials per run); nonetheless, across the whole experiment there were equal amount of trials for each combination. We also balanced the category used in the duplicate search task (butterfly, motorcycles and trees) across conditions and in combination with the category of the variable template (i.e. cow, dresser, skate). Each experimental run lasted ~7 minutes. The total duration of a session was ~1.7 hours (including, short brakes in between runs, the structural scan and mapper run).

9

The stimuli were back-projected on a 61 × 36 cm LCD screen (1920 x 1080 pixels) using Presentation (Neurobehavioral Systems, Albany, CA, USA) and viewed through a mirror attached to the head coil. Eye tracking data (EyeLink 1000, SR Research, Canada) was monitored to ensure participants were awake and attending the stimuli.

Anatomical Regions of interest (ROIs)

We created anatomically defined masks for four brain regions that have been previously found to participate in the maintenance or manipulation of VWM contents. We included the three ROIs as used by Christophel et al. (2018), namely early Visual Cortex (VC), the Intra Parietal Sulcus (IPS), Frontal Eye Fields (FEF), plus an additional area, namely the lateral Prefrontal Cortex (lPFC) based on the studies by Lee et al (2013) and Harel et al (2014). For the first three ROI masks we used the probabilistic atlas of visual topographic areas in standard space (i.e., MNI) developed by Wang, Mruczek, Arcaro and Kastner (2015). In our study, early visual cortex combined the ventral and dorsal regions of V1, V2 and V3 while the Intra Parietal Sulcus comprised IPS0, IPS1, IPS2, IPS3, IPS4 and IPS5. The lateral Prefrontal Cortex (lPFC) was mapped from the T1 scans of each subject, using the automatic cortical parcellation tool (i.e., recon-all) of Freesurfer (http://surfer.nmr.mgh.harvard.edu/), specifically the label used was the rostral-Middle-Frontal cortex (rMFC) from the Desikan atlas (Desikan et al., 2006), as Lee et al (2013) found similar results for functionally and anatomically defined ROIs, with their anatomical ROI being taken from the same atlas. Left and right hemisphere ROIs were combined into single masks, but prior to that we applied a threshold of 20% probability that the voxels in the ROI belonged to that specific region. This was done to reduce the size of the masks. Three main reasons led us to use merged (i.e., left and right hemisphere) ROIs: First, given that in our experiment the memorized object was not presented lateralized, we did not expect to find differences across hemispheres derived from retinotopic maps. Second, even when using lateralized stimuli, previous VWM studies found reliable effects in both hemispheres within VC, FEF (Christophel et al., 2018; Ester et al., 2009), IPS (Lorenc et al., 2018) and lPFC (Lee et al., 2013), with no reported substantial hemisphere differences. Finally, for reasons of comparison, we also report classification performance for pFs, which was mapped individually for each subject on the basis of an independent functional mapper (see van Loon et al. 2018 for details).

fMRI acquisition

Scanning was done on a 3T Philips Achieva TX MRI scanner with a 32-elements head coil. In the middle of the testing session (after four runs) a high-resolution 3D T1-weighted anatomical image (TR, 8.35 ms; TE, 3.83 ms; FOV, 240 × 220 × 188, 1 mm3 voxel size) was recorded for every

participant (duration 6 min). Blood oxygenation level dependent (BOLD)-MRI was recorded using Echo Planar Imaging (EPI) (TR 2000 ms, TE 27.62 ms, FA 76.1, 36 slices with ascending acquisition, voxel size 3 mm3, slice gap 0.3 mm, FOV 240 × 118.5×240).

fMRI data analysis

fMRI preprocessing. MRI data were registered to the subject-specific T1 scan using boundary-based registration (Greve & Fischl, 2009). The subject-specific T1 scan was registered to the MNI brain using FMRIB's Nonlinear Image Registration Tool (FNIRT). For the functional imaging data, we used FEAT version 5, part of FSL (Oxford Centre for Functional MRI of the Brain (FMRIB) Software Library; www.fmrib.ox.ac.uk/fsl; [[Smith et al., 2004]). Preprocessing steps consisted of motion correction, brain extraction, slice-time correction, alignment, and high-pass filtering (cutoff 100 s). For each subject and each run, two general linear models (GLM) were fitted to the data, whereby every TR (2 seconds each) was taken as a regression variable. In the first model, trials from all conditions (Current, Prospective and Irrelevant) contributed the same amount of regressors. Thus, this model contained 11 regressors as there were 11 TRs from the presentation of the memory item to one TR prior to the second search. The second model contained regressors from trials of the Current and Prospective condition only; specifically, from the onset of the second search display until the end of the trial. Therefore, this second model contained five regressors per trial per condition. We derived the t-value of each voxel for each of the 16 TRs in the experiment. We used FMRIB's Improved Linear Model (FILM) (Woolrich, Ripley, Brady, & Smith, 2001) for the time-series statistical analysis. The data were further analyzed in MATLAB (The MathWorks, Natick, MA, USA). For each participant, run, experimental condition (Current, Prospective and Irrelevant), category exemplar (Cow, Dresser and Skate, 4 exemplars of each) and TR, we created a vector containing the t-value per voxel in our four regions of interest. T-values for each predictor were computed by dividing the beta-weight by the standard error. That vector comprised the spatial pattern of activity evoked at that time point (TR) for that experimental condition in our region of interest.

Category decoding (Within-relevance and Cross-relevance) and Status decoding. We analyzed the multi-voxel patterns in two different ways: First, we performed category decoding to determine which ROIs carry information about the identity of the memorized objects and whether the neural representations of these objects varied with relevance (current, prospective or irrelevant). Second, we decoded the status of the memorized object to investigate which ROIs code for the relevance per se, regardless of object category. For this purpose we employed the Princeton Multi-Voxel Pattern Analysis toolbox (available at https://github.com/princetonuniversity/princeton-mvpa-toolbox, see Detre et al., 2006). When performing category decoding, for each condition and TR, a single class logistic regression classifier was trained to distinguish each object category (cow, dresser and skate). For status decoding, the classifiers learned to differentiate between conditions (Current vs. Prospective, Current vs. Irrelevant, and Prospective vs. Irrelevant) irrespective of object category. Logistic regression computes a weighted combination of voxel activity values, and it adjusts the (per-voxel) regression weights to minimize the discrepancy between the predicted output value and the correct output value. The maximum number of iterations used by the iteratively-reweighted least squares (IRLS) algorithm was set to 5000.

We performed the standard leave one run out cross validation procedure, which involved training a single class logistic regression classifier to learn a mapping between the neural patterns and the corresponding labels (i.e., either the categories [in the Category decoding] or the conditions [in the Status decoding]) for all but one run. Then, we used the trained classifier to predict the correct label (either category or condition) from the test patterns in the remaining run. For each iteration, we trained the classifier on eight runs and tested on the remaining run. This was done separately for each ROI, and each time point (TR) in the trial. This standard leave one out procedure was applied when doing relevance decoding and overall classification accuracy was the average accuracy of the nine iterations. Notice however that we could not fully balance the condition by object category combinations within runs (see Task and stimuli). Therefore, when performing category classification with the cross-validation procedure, each training set consisted of 8 runs with 32 trials per relevance condition which is not a multiple of 3 (i.e., the amount object categories of interest). Thus, for each relevance condition (Current, Prospective Irrelevant) when selecting all but one run for the training set, one of the categories contained 10 exemplars while the other two categories had 11. Likewise, the testing set contained all three categories (i.e., cow, dresser, skate), but two of the four exemplars belonged to the category that was less frequent in the training set. To make sure that this slight unbalance did not bias the classifiers against the least frequent category, we performed the leave one run out cross-validation procedure twice: once using the slightly unbalanced training set and another where we randomly picked one exemplar (for each of the

remaining two categories that had 11) and excluded them from the training set, leaving a balanced training set with 30 exemplars per relevance condition (10 per category). For both the balanced and unbalanced cross-validation procedures, overall classification accuracy was the average of the nine iterations. Finally, we used a balanced accuracy calculation as described in Fahrenfort, van Driel, van Gaal, & Olivers (2018) where accuracy is calculated separated per class and then averaged across classes. Because the results from both procedures were virtually the same, we chose to report analyses from the unbalanced procedure, where all available trials were included in the training sets.

We ran two types of category decoding: within the Current, Prospective and Irrelevant conditions (within-relevance decoding) and between relevance conditions (cross-relevance decoding). In the *within-relevance* decoding, we trained and tested classifiers within the same condition (Current, Prospective or Irrelevant) to determine if a given brain region contained any information about the category of the memorized item and if the strength of the neural representation varied across conditions. In the *cross-relevance* classification, we trained when the category was a current item and tested when the category was a prospective item and vice versa (Current-Prospective), to assess if category information was stored in the same format (predicting above-chance classification), the opposite format (predicting below-chance classification), or uncorrelated formats (predicting no correlation and thus classification at chance). We also included Current-Irrelevant and Prospective-Irrelevant comparisons. This resulted in six different testing and training combinations. To reduce the amount of comparisons, we averaged the classification performance of those combinations where the same conditions were used either for testing or training. Note that chance decoding for the category classifications was 33.33% since we had three object categories (Cow, Dresser and Skate).

For the *status* decoding, again for each brain region, condition and TR, a single class logistic regression classifier was trained, but now to distinguish between two conditions (for example current vs. prospective) rather than three categories. This resulted in three relevance decoding schemes: Current vs. Prospective, Current vs. Irrelevant and Prospective vs. Irrelevant. Because each classifier differentiated between two conditions only, chance decoding here was 50%.

We focused our statistical analyses on the average classification performance for each of three intervals in the trial, as predetermined on the basis of Lee, Kravitz, and Baker (2013) and our previous work (van Loon et al., 2018). Classification performance was then the average performance within the second, third and fourth TR after the offset of the cue (referred to as *Delay* interval), after the first search display (referred to as *Search 1* interval) or after the second search display (referred to as *Search 2* interval). The full time courses (classification performance per TR) are available in the supplementary material. The distribution of decoding scores revealed a significant skewness of .241

from normality (p < 0.05), so we first square-rooted the data (x' = √x) before running statistical tests. This reduced the skewness to a non-significant 0.062. We point out that the pattern of results also held for the non-transformed data. We compared category decoding performance and status decoding performance to their respective chance levels (33.33% and 50%) using two-tailed one-sample t-tests (following Christophel et al., 2018). We compared decoding performance across ROIs and relevance conditions using repeated measures ANOVA. All statistical analyses were performed using SPSS 26 (IBM, Armonk, USA), with α set at 0.05.

Results

We refer to Van Loon et al. (2018) for detailed behavioural results. In short, performance in the template-based search task and the duplicate search task was comparable in terms of accuracy, though somewhat slower on the latter. Overall, template-based search was faster in the first than in the second search task, while the speed of duplicate search did not differ between first and second search.

Our fMRI analyses targeted four ROIs: Visual Cortex (VC), Intra Parietal Sulcus (IPS), Frontal Eye Fields (FEF) and lateral Prefrontal Cortex (lPFC). To investigate how category representation changed with relevance (i.e., the condition: Current, Prospective, and Irrelevant) across these ROIs, we trained classifiers on the multivoxel response patterns in each of those areas, separately for each repetition time (TR). First, we trained and tested the classifier with trials in which the memorized category was currently relevant (for Search 1), when it was prospectively relevant (for Search 2) and when it was not a target in the trial (i.e. irrelevant condition). This *within*-relevance decoding scheme evaluates if any given brain area codes for the category or identity of the memorized items when current, prospective or irrelevant, but it does not reveal whether these category representations differ across the three levels of relevance. Therefore, we additionally planned a *cross*-relevance decoding scheme in which we trained the classifier when the objects were currently relevant and tested when the same objects were prospectively relevant and vice versa (Current-Prospective cross-relevance scheme). We did the same for the other two possible combinations: Current-Irrelevant and Prospective-Irrelevant. Crucially, if the category representations in two of the conditions (e.g., Current and Prospective) are similar, above-chance classification accuracy is expected again here. If representations are dissimilar in an unrelated fashion, classification is expected to be at chance levels, while below-chance classification is predicted when representations are dissimilar, but in a systematic, anti-correlated fashion.

Finally, we ran a status decoding analysis, to identify those brain areas that preferentially code for the behavioral relevance of the memorized item, irrespective of object category. To this end, we trained and tested the classifier on the distinction between Current and Prospective conditions, Current and Irrelevant, and Prospective and Irrelevant. Here we focus on three pre-determined intervals, referred to as Delay, Search 1, and Search 2 (see Methods). Supplementary Figures S1, S2 and S3 show the full time course of the within-relevance category decoding, cross-relevance category decoding and the status decoding respectively, for each TR. The time courses for pFs were reported in van Loon et al (2018). Note that especially the Delay interval – that is prior to the first search – is of interest here, since any status effects during the searches likely reflect that participants were doing a particular search task.

The delay prior to the first search: Posterior brain areas preferentially code for the content, while frontal areas only code for the relevance of the memory item.

Figure 2 shows the decoding accuracy for each type of classification analysis, ROI and condition during the Delay interval prior to search. First, using the within-relevance classification scheme, we evaluated which brain areas carry information about the category of the memorized item (i.e., cow, dresser or skate) and if the strength of the category representation varies with its behavioral relevance. As can be seen in Figure 2A, we could reliably decode the category of the memorized item in VC, pFs and IPS, but not in FEF or lPFC (See Supplementary Table 1 for all one sample t-tests against chance, 33.33%). To assess whether category information varies for different regions of interest, we ran a two-way repeated measures ANOVA on the same data, with factors ROI (VC, pFs, IPS, FEF, lPFC) and Condition (Current, Prospective and Irrelevant). There was a main effect of Condition on category decoding ($F_{(2,48)}$ = 3.48, p = 0.039, η_p^2 = 0.13), as category decoding was better for currently and prospectively relevant objects than for irrelevant objects. There was also a main effect of ROI ($F_{(4,96)}$ = 24.82, p < 0.001, η_p^2 = 0.51), as category decoding decreased further up the cortical hierarchy for parietal and frontal control structures compared to sensory areas (VC and pFS). The interaction between ROI and Condition failed to reach significance (F = 1.79, p = 0.08). As previous studies suggest that the strength of VWM representations may not be equal for all IPS sub regions, supplementary Figure S4 shows the classification analysis for each of the IPS sub region separately from IPS0 to 5. Results revealed that the content representation is stronger for the lower IPS areas (specifically IPS0, IPS1 and IPS2), consistent with earlier findings (Bettencourt & Xu, 2016; Xu & Chun, 2006).

Next, we used the cross-relevance decoding scheme to assess whether current, prospective and irrelevant items of the same category shared a similar neural representational pattern. Figure 2B shows that the posterior areas –VC, pFs and IPS – exhibited above chance classification for all condition combinations, whether Current-Prospective, Current-Irrelevant, or Prospective-Irrelevant, indicating that during the delay interval prior to search, the neural representation of the same category was very similarin those brain areas, regardless of its current, prospective or irrelevant status (see also Supplementary Table 1 for all comparisons against chance, 33.33%, and see Supplementary Figure S4 for the individual IPS regions). The frontal areas again showed weaker evidence for category information with some evidence for common coding from the Current-Irrelevant decoding scheme in FEF and Current-Prospective decoding scheme in lPFC. A two-way repeated measures ANOVA with ROI and relevance condition as factors revealed a main effect of ROI ($F_{(4,96)}$ = 18.81, p < 0.001, η_p^2 = 0.44) confirming that the similarity in the category patterns across conditions decreased for fronto-parietal brain regions. We also observed a main effect of Decoding scheme ($F_{(4,96)}$ = 5.01, p = 0.011, η_p^2 = 0.17), as cross-decoding was stronger for current and prospective representations than for either type of representations paired with irrelevant representations. There was no interaction between the two factors F = .6, p = .75). Thus overall, during the delay period, the pattern across different relevance conditions was very much the same as for within-relevance decoding.

Finally, we used the *status* decoding scheme to assess which brain areas code for the behavioral relevance of the memorized item regardless of its category. We ran status decoding separately for each pair of conditions (Current-Prospective, Current-Irrelevant, Prospective-Irrelevant). As can be seen in Figure 2C, we could decode the status of the memorized item in all ROIs (see also Supplementary Table 1 for all comparisons against chance, 50%, and see Supplementary Figure S4 for the individual IPS regions). A two-way repeated measures ANOVA on the status decoding accuracy with ROI and specific status decoding scheme as factors revealed a main effect of ROI ($F_{(4,96)}$ = 22.13, p < 0.001, η_p^2 = 0.48) with a significant linear relationship ($F_{(1,24)}$ = 39.04, p < 0.001, η_p^2 = 0.61). In contrast to category decoding, here status decoding increased for brain regions higher in the cortical hierarchy, specifically IPS, FEF, and lPFC. The main effect of the specific status decoding scheme was also significant ($F_{(4,96)}$ = 17.36, p < 0.001, η_p^2 = 0.42). Across most ROIs, the strongest distinction was between currently relevant and irrelevant, followed by between prospectively relevant and irrelevant. Importantly though, currently relevant and prospectively relevant could also be clearlydistinguished, especially in IPS and lPFC. The exception was VC, which could not discriminate between current and prospective status, which resulted in a significant ROI by Decoding scheme interaction ($F_{(8,192)}$ = 2.47, p = 0.015, η_p^2 = 0.09).

These results indicate that during the delay prior to search, the posterior sensory related brain areas such VC and pFs carry mainly information about the category of the memorized item, while the frontal areas FEF and lPFC are more strongly involved representing the status of the item, while IPS revealed both category decoding and strong status decoding. Figure 3 summarizes this finding as it directly juxtaposes category and relevance decoding performances. An ANOVA with decoding scheme (category decoding versus relevance decoding) and ROI as factors statistically confirmed the interaction that is evident from the graph ($F_{(4,96)}$ = 36.35, $p < 0.001$, η_p^2 = 0.60) and supports the opposite trends.

A.

Category Decoding (Within- Relevance)

B.

Category Decoding (Cross-Relevance)

C.

Relevance Decoding

Figure 2. Delay interval: Object category decoding (Within-relevance and cross-relevance) and status decoding for VC, pFs, IPS, FEF and lPFC. Previously reported data from pFs were also included for comparison

purposes. **(A) Category classification performance within each relevance condition.** More posterior brain areas (VC, pFs and IPS) contained information about the identity (i.e., category specificity) of current, prospective and irrelevant items (the latter only in VC and pFs), whereas frontal ROIs (FEF and lPFC) did not. **(B) Category classification performance using cross-relevance decoding.** The more posterior brain areas (VC, pFs and IPS) showed above chance classification for all decoding combinations, suggesting similar representations for current, prospective and irrelevant objects. In FEF category representations of current and irrelevant items were similar, while all other decoding combination were at chance. In lPFC the category representation of current and prospective items were similar, while Current-Irrelevant and Prospective-Irrelevant representations were uncorrelated. **(C) Status decoding.** Status decoding increased for regions higher in the cortical hierarchy (IPS, FEF and lPFC). Except for VC, for all areas differentiating between currently relevant and completely irrelevant representations (Current-Irrelevant) was better than between currently relevant and prospectively relevant representations (Current-Prospective) or between prospectively relevant and irrelevant representations (Prospective-Irrelevant). Error bars indicate Standard Error (SE) of the mean, N= 25. * $p<0.05$, ** $p < .01$, *** $p< .001$.

Figure 3. Delay interval: Direct comparison of category decoding versus relevance (status) decoding. Here classification performance is shown collapsed across the three sub conditions (current, prospective, irrelevant) and with chance level subtracted (33.33% for category decoding, 50% for relevance decoding). Error bars indicate Standard Error of the mean (SE).

Search 1: Dissimilar representations for current and prospective items

Figure 4 shows the decoding accuracy for the Search 1 interval. As can be seen in Figure 4A, in most ROIs only the current category was decodable, while the other two conditions were at chance, the exception being pFs where we could also decode the category of the prospectively relevant item (See Supplementary Table 2 for comparisons against chance: 33.33% in the Search 1 interval). This greater classification accuracy for current representations was to be expected since during the first search, they were not only more relevant, but also presented on the screen. A two-way ANOVA on the within-relevance decoding accuracy showed a significant main effect of Condition ($F_{(2,48)}$ = 32.83, p < 0.001, η_p^2 =0.58), confirming the stronger decoding for the current condition. There was also a significant main effect of ROI ($F_{(4,96)}$ = 17.84 p < 0.001, η_p^2 = 0.43). As was the case for the delay period, overall category decoding was stronger for posterior than for frontal brain regions. Finally, we observed a significant interaction ($F_{(8,192)}$ = 2.87, p = 0.001, η_p^2 = 0.11), as the differences in category decoding between conditions were larger in pFs than in other ROIs.

Next, we used the cross-relevance decoding scheme to assess whether current, prospective and irrelevant items shared the same neural pattern. As shown in Figure 4B, a number of areas showed significant *below* chance decoding during the Search 1 interval (see Supplementary Table 2 for comparisons against chance: 33.33%). Most notably pFS (as already reported in Van Loon et al., 2018), but also areas FEF and lPFC showed reliable below chance decoding performance (with IPS and VC showing numerical effects in the same direction). A two-way repeated measures ANOVA revealed no effect of ROI (F = 1.6, p = .175). We found a main effect of Decoding scheme ($F_{(2, 48)}$ = 8.39, p = 0.001, η_p^2 = 0.26). On average across ROIs, current and prospective representations of the same category were more *dis*similar to each other than when comparing current to irrelevant or prospective to irrelevant representations. There was also an interaction effect ($F_{(8,192)}$ = 2.03, p= 0.044, η_p^2 = 0.078), which reflected the fact that these (dis)similarity effects were strongest in pFs. We did not perform any additional statistical analyses on the status decoding for this interval, as any status differences would only reflect the fact that observers actively performed a search task for one stimulus category (in the Current condition). For the same reason we did not directly juxtapose category versus relevance status decoding here. However, for completion, we included the average classification performance for each status classification scheme and ROI in Figure 4C.

In conclusion, the results confirm an overall decrease in category sensitivity towards more frontal areas. In addition, current and prospective items were represented in anti-correlated patterns (as indicated by the below-chance decoding) in multiple areas, most notably pFs, but also in frontal cortex.
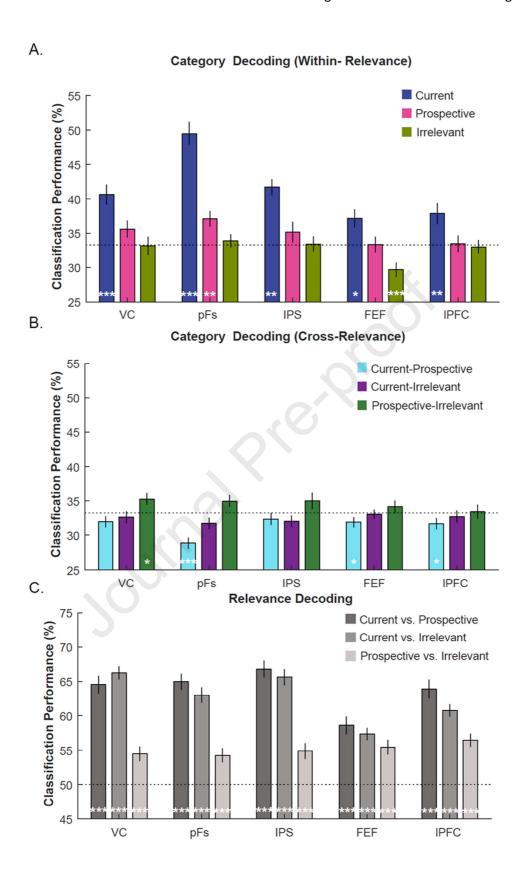
A.

Category Decoding (Within- Relevance)

B.

Category Decoding (Cross-Relevance)
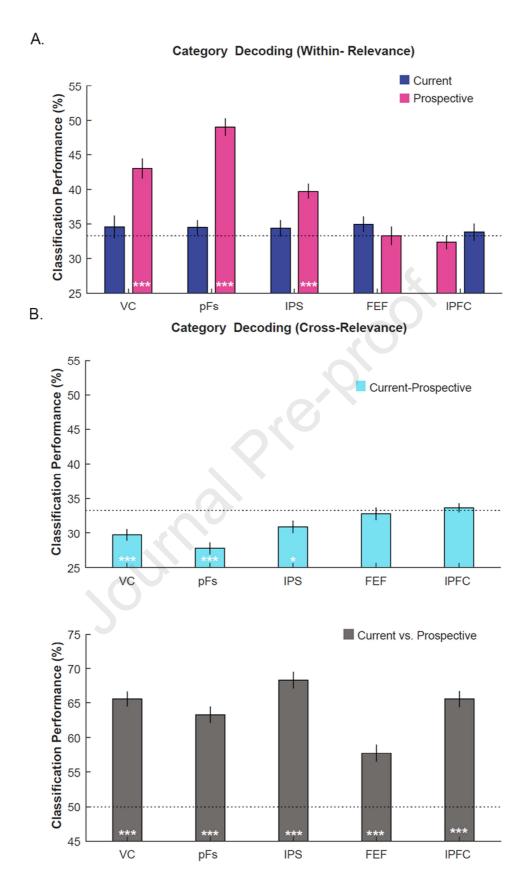

C.

Relevance Decoding

Figure 4. Search 1 interval: Object category decoding (Within-relevance and cross-relevance) for VC, pFs, IPS, FEF and lPFC. Data from pFs were also included for comparison purposes. **(A) Category classification performance for each relevance condition.** In all areas, category decoding was higher for the Current than for

the Prospective and Irrelevant conditions. We could reliably decode the category of prospectively relevant items in pFs and to a lesser extent VC. **(B) Category classification performance across relevance conditions.** Only in pFs we observed below chance decoding for the Current-Prospective decoding scheme, indicating that the category representation of prospective items is inverted relative to current representation of the same category. VC did not reliably show this effect, although the overall pattern was in the same direction, as was the case for most other areas. **(C) Status decoding.** For completion, we show status decoding. Notice however that here decoding accuracy is likely to be driven by the differences between the two types of search task (Template search vs Duplicate search) rather than the relevance of the memorized object (Current vs Prospective or Current vs Irrelevant). The only meaningful classification scheme here is Prospective vs. Irrelevant (light grey bars): in these two conditions participants performed the same type of task as first search (i.e., the duplicate search), and still, the classifier could discriminate between trials where the item in memory was only temporarily irrelevant (prospective condition) from trials where the item was completely irrelevant. Error bars indicate Standard Error of the mean (SE), N= 25. * p<0.05, **p < .01, ***p< .001.

Search 2: Inverse representations of the previous target in VC, pFs and IPS

Figure 5 shows the decoding accuracy for each type of classification analysis, ROI and condition in the Search 2 interval. Note that there is no data for the Irrelevant condition here as in that condition there was no second search. As can be seen in Figure 5A, during the Search 2 interval, only prospectively relevant items were decodable in VC, pFs, and IPS (See Supplementary Table 3 for comparisons against chance: 33.33%). This greater decoding accuracy for prospective than current items was to be expected since during the second search of the Prospective condition, the prospective category was actually to be searched and thus on the screen. A two-way repeated measures ANOVA on the within-relevance decoding accuracy revealed a significant main effect of relevance condition ($F_{(2,48)}$ = 25.49, p < 0.001, η_p^2 = 0.52). There was also a main effect of ROI ($F_{(4,96)}$ = 16.09 p < 0.001, η_p^2 = 0.40), as well as a significant interaction between the two factors ($F_{(8,192)}$ = 13.76, p < 0.001, η_p^2 = 0.36). Similar to what we observed in the Delay and Search 1 intervals, during the Search 2 interval, category decoding decreased for brain regions higher in the cortical hierarchy, here for the prospective category. While the more posterior brain areas (i.e., VC, pFs and IPS) were sensitive to the prospective category information, the frontal ROIs (FEF and lPFC) were not. In contrast, for the now abandoned (current) search target, there was little evidence in any ROI.

Next, we used the cross-relevance decoding scheme to assess whether the prospective target and the now abandoned target shared the same or different representational patterns. As can be seen in Figure 5B, here VC, pFs and IPS showed below-chance decoding during the Search 2 interval (see Supplementary Table 3 for comparisons against chance: 33.33%), while frontal areas did

not. Accordingly, a one-way repeated measures ANOVA on the Current-Prospective cross-relevance scheme revealed a significant main effect of ROI ($F_{(4,96)}$ = 11.8, p < 0.001, η_p^2 = 0.33). Again, we did not perform any additional statistical analyses on the status decoding for this interval, as this would only reflect the ongoing task. However, for completion, we included the average classification performance for each status classification scheme and ROI is Figure 5C. As can be seen, the activation of the prospective task can be picked up across all ROIs.

In summary, during the Search 2 interval, we observed stronger category decoding for prospective than for the now abandoned "current" items. Moreover, posterior brain areas such as VC and pFs but also IPS showed below-chance cross-relevance decoding, indicating a dissimilarity between abandoned target representations and reactivated target representations.

A.

B.

Figure 5. Search 2 interval: Object category decoding (Within-relevance and cross-relevance) for VC, pFs, IPS, FEF and IPFC. Data from pFs were also included for comparison purposes. (A) Category classification performance within-relevance. In VC, pFs and IPS we could decode the category of the prospective items, with

higher accuracy in the Prospective than in the Current condition, while FEF and lPFC did not contained category information. **(B) Category classification performance cross-relevance.** The most posterior areas VC, pFs and IPS show below chance decoding for the Current-Prospective decoding scheme, indicating that the category representation of prospective items was inverted relative to current representations. Note that there was no irrelevant condition for the second search. **(C) Status decoding.** For completion, we show status decoding. Notice however that here the decoding accuracy is likely to be driven by the differences between the two types of search task (Template search vs Duplicate search) rather than the relevance of the memorized object (Current vs Prospective). Error bars indicate Standard Error (SE), N= 25. * p<0.05, **p < .01, ***p< .001.

Discussion

Previous studies have shown that multiple brain regions participate in the process of representing and manipulating working memory content (Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017; D'Esposito & Postle, 2015; Postle, 2006; Ranganath et al., 2004). Moreover, working memory representations seem to flexibly adapt according to task demands, in terms of strength (Larocque et al., 2014; LaRocque, Postle, Riggall, & Emrich, 2016; Lewis-Peacock et al., 2012), representational content (Lee et al., 2013; Lewis-Peacock, Drysdale, & Postle, 2015; Yu & Shim, 2017) or neural pattern (van Loon et al., 2018) of the memorized objects. In the current study we assessed the role of posterior (i.e. VC, pFs and IPS) as well as frontal (FEF and lPFC) cortices in the representation of working memory items as they are required in a sequence of multiple tasks. The goal was to investigate how the representation of both the category and the relevance of the VWM content varies along the cortical hierarchy. In particular, we were interested in which brain areas are sensitive to VWM *content*, and which areas are sensitive to the *status* of the memory – that is whether the memory is relevant for current versus prospective task goals.

Our results revealed a clear division of work across brain regions: Posterior brain areas preferentially coded for the content (i.e. category) of the memorized items, while frontal areas mostly carried information about the relevance status of items in working memory. During the delay period prior to search, we could recover the category of the memorized item from VC and IPS, with stronger category representation when the object was behaviorally relevant (either current or prospective) than when completely irrelevant. As shown earlier (van Loon et al., 2018), pFS also contains category information, as would be expected. Areas FEF and lPFC on the other hand showed little to no category specificity during the delay, and therefore also no clear status-related modulation of such category-specific activity. The opposite trend emerged when evaluating which

brain regions are involved in the representation of the status of the memory, *irrespective* of object category – that is whether an item was currently relevant, prospectively relevant, or not relevant at all. Although we could decode the memory status from virtually all areas, including VC and pFs, relevance discrimination became considerably stronger for more anterior regions, notably IPS, FEF and lPFC. Note that the opposite trends preclude an explanation in terms of mere differences in signal to noise levels across regions, and must thus represent different functionalities.

Dominant views propose that working memory emerges from the flexible contribution of multiple brain regions with sensory areas representing sensory-based content information while prefrontal regions encode more abstract, semantic, and response-related aspects of memorized stimuli. While so far we have rather loosely and intuitively referred to the posterior-anterior cortical hierarchy, the crucial question remains what the exact functional roles are of the areas involved. In our view a reasonably clear picture of prefrontal functioning is emerging, as it provides the goal- or task context-related signals that prioritize relevant over irrelevant representations in working memory (e.g. Christophel et al., 2017; Fuster, Bauer, & Jervey, 1985; Lee & D'Esposito, 2012; Miller, Vytlacil, Fegen, Pradhan, & D'Esposito, 2010; Serences, 2016; Zanto, Rubens, Thangavel, & Gazzaley, 2011). Consistent with this, we found PFC and FEF regions to predominantly code for task relevance, a more abstract property than stimulus content (although the latter was not completely absent). In contrast, the exact functional role(s) of the posterior parietal cortex in working memory remains largely a puzzle (Postle, 2020; Xu, 2018; 2020).

One hypothesis has been that IPS representations serve as a more robust backup copy of sensory content than the sensory representations themselves, protecting memories against interference (Bettencourt & Xu, 2016; Christophel et al., 2018; Lorenc et al., 2018). This parietal representation may then in addition allow for a template matching process at the memory test, the outcome of which is then relayed to frontal regions to presumably interact with task and response settings (Xu, 2020). However, as Postle and Yu (2020) have recently warned, the fact that one can discriminate memory content in certain brain regions does not necessarily mean that these regions really represent memory content. Rather, the posterior parietal cortex could be serving only control operations, with the specific parameters of these operations depending on the stimulus. Here in our study we found that IPS codes for relevance as well as for stimulus category. At a general level of description it is therefore tempting to see posterior parietal cortex (here sepecifically IPS) as a highly flexible hub that multiplexes task rules with stimulus content, to create a coherent representation or priority map of which stimulus or memory should currently guide behavior (Cole et al., 2013; Majerus et al., 2007; Mitchell & Cusack, 2007; Silver, Ress, & Heeger, 2005; Sprague, Itthipuripat, Vo,

& Serences, 2018; Sprague & Serences, 2013; Xu, 2007). Xu (2018) for example has raised the hypothesis that while the temporal lobe serves to provide as rich as possible knowledge of the world, parietal cortex serves the opposite, namely to make a selection in order to guide behavior (which is inherently limited). Parietal cortex thus extracts the parameters to support the execution of appropriate actions (cf. Cisek, 2007). In line with such a multiplexing role between sensory and frontal cortex, there is recent evidence for the binding of stimulus information to order and context within working memory (e.g. Yu and Shim 2019; Gosseries et al. 2018).

We found little evidence supporting the idea that in anticipation of, and during, the first task, content-specific information for the prospective task was transformed from early visual representations towards higher cortical representations instead (most notably IPS or FEF representations; Bettencourt & Xu 2016; Christophel et al 2018). We found no evidence for prospective representations being actively represented in IPS and FEF during the first search task, while category information on the search target itself could be decoded from these areas, indicating that these areas could in principle hold category information. Together, these findings are in contrast to those by Christophel et al. (2018), who could classify specific orientations of both currently and prospectively relevant representations from FEF. Of course, the absence of prospective item-related activity in FEF in our study does not preclude the possibility that FEF contains stimuli-specific information, nor does it exclude the possibility of a partial transfer in representation from lower sensory areas to frontal areas. Our study may have been underpowered to detect a real but relatively small effect. Indeed, our sample size was less than a third of that of Christophel and colleagues, who tested an exceptionally large and laudable number of participants (89 subjects, vs. 25 here). In addition, our selection of ROIs may not have been optimal for detecting potentially more subtle multivariate patterns in subregions. However, a distinct alternative possibility is that earlier-found decoding of orientation in FEF does not reflect orientation representations *per se*, but instead is caused by participants' systematic eye movements. Consistent with this idea, Mostert et al. (2018) showed that the decoding analyses of stimulus orientation in a VWM task were heavily confounded by the eye movements that participants made during the delay period, which were systematically related to the identity of the memorized stimulus. Note that this concern is especially valid for FEF, which is directly involved in eye movement planning (Paus, 1996; Schall & Hanes, 1993; Schiller, True, & Conway, 1980; Vernet, Quentin, Chanes, Mitsumasu, & Valero-Cabré, 2014).

It is worth pointing out another aspect of the results with respect to the questions whether frontal areas contain category information. Note that while the target category could be reliably

decoded from frontal areas during the first search, this was not the case for the second search. At the second search, the up to then prospective target becomes the new "current" search target (although in our graphs it retains the label prospective). Yet while relevant for the task, the second search target could not be reconstructed from FEF and lPFC activity patterns. This was not due to second targets being represented more weakly overall, as in posterior areas it could be recovered at similar levels as for the first search. One possibility is that these frontal areas become more heavily involved in biasing category information if there is still the potential for competition between multiple tasks, as is the case during the first search. During the second search only one task remains, and there may thus be less need to protect or bias its target representation.

Finally, we assessed whether those areas that represented memory content did so similarly for currently and prospectively relevant items. Using a cross-relevance decoding scheme (where the classifier was trained to discriminate between categories for one status but tested for the other status), we observed above-chance classification performance during the delay in VC, pFS, and IPS, indicating that in those areas the neural representation of the object categories was very similar regardless of its current or prospective status. In our previous report (van Loon et al., 2018), when using the same cross-relevance decoding scheme, we observed a transformation from above-chance decoding during the delay to below-chance decoding during the first search interval within the pFs area. We hypothesized that this may reflect a mechanism to shield representations from interference with the ongoing task, through suppression of the associated neuronal population. Here we assessed if such inversion also occurred in other brain regions throughout the cortical hierarchy. Although numerically such inversions were present in all ROIs, they were reliable only in FEF and lPFC. Thus, it appears that FEF and lPFC represented the content of prospective memory at some level, but in a pattern opposite to when the target object is active. We speculate that observers may temporarily suppress the prospective representation in order to prevent interference with the ongoing task, and that this suppression may originate in frontal areas. Such pattern inversion or suppression was rather weak in early visual cortex. We also speculate that pattern inversion will mainly occur in regions that specialize in representing the memory content, as to optimally prevent interference. Consistent with this idea, Yu, Teng, and Postle (2020) have recently reported evidence for rotated representational patterns in VC for a different class of stimulus, for which VC is highly specialized, namely oriented gratings. Notice that we did find more robust below-chance decoding during the second search interval, for the by then no longer relevant first search target, in both VC and IPS – again similar to what we previously observed in pFs (van Loon et al., 2018). This may reflect the active purging of a no-longer relevant search target, or more passive remnants of no longer relevant representations. More research is required to elucidate under which circumstances

the pattern inversion occurs as well as its underlying mechanism. What can be concluded though is that during the search tasks, current and prospective items were dissimilar from each other across a wide range of brain areas.

In sum, we provide further evidence for the distributed nature of visual working memory processes by showing a division of work across brain regions. While posterior sensory cortices mainly represent stimulus content, frontal cortex codes for the current or prospective status of the memory. IPS is then the hub that connects content and task relevance, setting current stimulus priorities. Furthermore, current and prospective representations may be kept apart by anti-correlating their representations.

Acknowledgments

References

Bettencourt, K. C., & Xu, Y. (2015). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience*, *19*(1), 150–157. http://doi.org/10.1038/nn.4174

Christophel, T. B., Hebart, M. N., & Haynes, J.-D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(38), 12983–9. http://doi.org/10.1523/JNEUROSCI.0184-12.2012

Christophel, T. B., Iamshchinina, P., Yan, C., Allefeld, C., & Haynes, J. D. (2018). Cortical specialization for attended versus unattended working memory. *Nature Neuroscience*, *21*(4), 494–496. http://doi.org/10.1038/s41593-018-0094-4

Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J. D. (2017). The Distributed Nature of Working Memory. *Trends in Cognitive Sciences*, *21*(2), 111–124. http://doi.org/10.1016/j.tics.2016.12.007

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis.

Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1485), 1585-1599.

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., & Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature Neuroscience*, *16*, 1348. Retrieved from https://doi.org/10.1038/nn.3470

Conde-Sousa, E., & Aguiar, P. (2013). A working memory model for serial order that stores information in the intrinsic excitability properties of neurons. *Journal of Computational Neuroscience*, *35*(2), 187–199. http://doi.org/10.1007/s10827-013-0447-7

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology*, *66*, 115–142. http://doi.org/10.1146/annurev-psych-010814-015031

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. http://doi.org/https://doi.org/10.1016/j.neuroimage.2006.01.021

Detre, G., Polyn, S. M., Moore, C. D., V.S., N., Singer, B., Cohen, J. ., & Norman, K. A. (2006). Princeton multi-voxel pattern analysis (MVPA) toolbox. In *In Poster Presented at the Annual Meeting of the Organization for Human Brain Mapping*.

Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed Patterns of Activity in Sensory Cortex Reflect the Precision of Multiple Items Maintained in Visual Short-Term Memory. *Journal of Neuroscience*, *33*(15), 6516–6523. http://doi.org/10.1523/JNEUROSCI.5732-12.2013

Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially Global Representations in Human Primary Visual Cortex during Working Memory Maintenance. *Journal of Neuroscience*, *29*(48), 15258–15265. http://doi.org/10.1523/JNEUROSCI.4388-09.2009

Fahrenfort, J. J., van Driel, J., van Gaal, S., & Olivers, C. N. L. (2018). From ERPs to MVPA using the Amsterdam Decoding and Modeling toolbox (ADAM). *Frontiers in Neuroscience*, *12*. http://doi.org/10.3389/fnins.2018.00368

Fuster, J. M., Bauer, R. H., & Jervey, J. P. (1985). Functional interactions between inferotemporal and prefrontal cortex in a cognitive task. *Brain Research*, *330*(2), 299–307. http://doi.org/10.1016/0006-8993(85)90689-4

Gosseries, O., Yu, Q., LaRocque, J. J., Starrett, M. J., Rose, N. S., Cowan, N., & Postle, B. R. (2018).

Parietal-occipital interactions underlying control-and representation-related processes in
working memory for nonspatial visual features. *Journal of Neuroscience*, *38*(18), 4357-4366.

Greve, D. N., & Fischl, B. (2009). Accurate and Robust Brain Image Alignment using Boundary- based
Registration. *Neuroimage*, *48*(1), 63–72.
http://doi.org/10.1016/j.neuroimage.2009.06.060.Accurate

Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing
differentially across the cortex. *Proceedings of the National Academy of Sciences of the United
States of America*, *111*(10), E962–E971. http://doi.org/10.1073/pnas.1312567111

Harrison, S. a, & Tong, F. (2009). Decoding reveals the contents of visual working memory in early
visual areas. *Nature*, *458*(7238), 632–5. http://doi.org/10.1038/nature07832

Larocque, J. J., Lewis-Peacock, J. a, & Postle, B. R. (2014). Multiple neural states of representation in
short-term memory? It's a matter of attention. *Frontiers in Human Neuroscience*, *8*, 5.
http://doi.org/10.3389/fnhum.2014.00005

LaRocque, J. J., Postle, B. R., Riggall, A. C., & Emrich, S. M. (2016). Within-Category Decoding of
Information in Different Attentional States in Short-Term Memory. *Cerebral Cortex*, *27*(10),
4881–4890. http://doi.org/10.1093/cercor/bhw283

Larocque, J. J., Riggall, A. C., Emrich, S. M., & Postle, B. R. (2017). Within-Category Decoding of
Information in Different Attentional States in Short-Term Memory. *Cerebral Cortex*, *27*(10),
4881–4890. http://doi.org/10.1093/cercor/bhw283

Lee, S., Kravitz, D. J., & Baker, C. I. (2013). Goal-dependent dissociation of visual and prefrontal
cortices during working memory. *Nature Neurosci.*, *16*(8), 997–999.
http://doi.org/10.1038/nn.3452.Goal-dependent

Lee, T. G., & D'Esposito, M. (2012). The Dynamic Nature of Top-Down Signals Originating from
Prefrontal Cortex: A Combined fMRI-TMS Study. *Journal of Neuroscience*, *32*(44), 15458–15466.
http://doi.org/10.1523/jneurosci.0627-12.2012

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*(1), 61–79. http://doi.org/10.1162/jocn_a_00140

Lewis-Peacock, J. A., Drysdale, A. T., & Postle, B. R. (2015). Neural evidence for the flexible control of
mental representations. *Cerebral Cortex*, *25*(10), 3303–3313.

http://doi.org/10.1093/cercor/bhu130

Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R. E., & D'Esposito, M. (2018). Flexible coding of visual working memory representations during distraction. *The Journal of Neuroscience*, *38*(23), 3061–17. http://doi.org/10.1523/JNEUROSCI.3061-17.2018

Majerus, S., Bastin, C., Poncelet, M., Van der Linden, M., Salmon, E., Collette, F., & Maquet, P. (2007). Short-term memory and the left intraparietal sulcus: Focus of attention? Further evidence from a face short-term memory paradigm. *NeuroImage*, *35*(1), 353–367. http://doi.org/10.1016/J.NEUROIMAGE.2006.12.008

Mansouri, F. A., Koechlin, E., Rosa, M. G. P., & Buckley, M. J. (2017). Managing competing goals — a key role for the frontopolar cortex. *Nature Reviews Neuroscience*, *18*, 645. Retrieved from https://doi.org/10.1038/nrn.2017.111

Miller, B. T., Vytlacil, J., Fegen, D., Pradhan, S., & D'Esposito, M. (2010). The Prefrontal Cortex Modulates Category Selectivity in Human Extrastriate Cortex. *Journal of Cognitive Neuroscience*, *23*(1), 1–10. http://doi.org/10.1162/jocn.2010.21516

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural Mechanisms of Visual Working Memory in Prefrontal Cortex of the Macaque. *The Journal of Neuroscience*, *16*(16), 5154–5167. http://doi.org/https://doi.org/10.1523/JNEUROSCI.16-16-05154.1996

Mitchell, D. J., & Cusack, R. (2007). Flexible, Capacity-Limited Activity of Posterior Parietal Cortex in Perceptual as well as Visual Short-Term Memory Tasks. *Cerebral Cortex*, *18*(8), 1788–1798. http://doi.org/10.1093/cercor/bhm205

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic Theory of Working Memory. *Science*, *1543*(March), 1543–1547. http://doi.org/10.1126/science.1150769

Mostert, P., Albers, A. M., Brinkman, L., Todorova, L., Kok, P., & de Lange, F. P. (2018). Eye Movement-Related Confounds in Neural Decoding of Visual Working Memory Representations. *Eneuro*, *5*(4), ENEURO.0401-17.2018. http://doi.org/10.1523/eneuro.0401-17.2018

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends in Cognitive Sciences*, *15*(7), 327–34. http://doi.org/10.1016/j.tics.2011.05.004

Parthasarathy, A., Herikstad, R., Bong, J. H., Medina, F. S., Libedinsky, C., & Yen, S. C. (2017). Mixed selectivity morphs population codes in prefrontal cortex. *Nature Neuroscience*, *20*(12), 1770–

1779. http://doi.org/10.1038/s41593-017-0003-2

Paus, T. (1996). Location and function of the human frontal eye-field: A selective review. *Neuropsychologia*, *34*(6), 475–483. http://doi.org/10.1016/0028-3932(95)00134-4

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. http://doi.org/10.1016/j.neuroscience.2005.06.005

Postle, B. R., Ferrarelli, F., Hamidi, M., Feredoes, E., Massimini, M., Peterson, M., … Tononi, G. (2006). Repetitive Transcranial Magnetic Stimulation Dissociates Working Memory Manipulation from Retention Functions in the Prefrontal , but not Posterior Parietal , Cortex. *Journal of Cognitive Neuroscience*, *18(10)*(1999), 1712–1722. http://doi.org/https://doi.org/10.1162/jocn.2006.18.10.1712

Postle, B. & Yu, Q. (2020). Neuroimaging and the localization of function in visual cognition. *Visual Cognition*

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2018). Simultaneous representation of sensory and mnemonic information in human visual cortex. *BioRxiv Preprint*. http://doi.org/doi: http://dx.doi.org/10.1101/339200

Rainer, G., Asaad, W. F., & Miller, E. K. (1998). Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature*, *393*(6685), 577–579. http://doi.org/10.1038/31235

Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior Temporal, Prefrontal, and Hippocampal Contributions to Visual Working Memory Maintenance and Associative Memory Retrieval. *Journal of Neuroscience*, *24*(16), 3917–3925. http://doi.org/10.1523/JNEUROSCI.5053-03.2004

Riggall, A. C., & Postle, B. R. (2012). The Relationship between Working Memory Storage and Elevated Activity as Measured with Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *32*(38), 12990–12998. http://doi.org/10.1523/JNEUROSCI.1892-12.2012

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*, 585. Retrieved from https://doi.org/10.1038/nature12160

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., … The. (2016).

Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, *354*(6316), 1136–1139. http://doi.org/10.1038/nrn.2017.113

Schall, J. D., & Hanes, D. P. (1993). Neural basis of saccade target selection in frontal eye field during visual search. *Nature*, *366*(6454), 467–469. http://doi.org/10.1038/366467a0

Schiller, P. H., True, S. D., & Conway, J. L. (1980). Deficits in eye movements following frontal eye-field and superior colliculus ablations. *Journal of Neurophysiology*, *44*(6), 1175–1189. http://doi.org/10.1152/jn.1980.44.6.1175

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, *128*, 53–67. http://doi.org/10.1016/J.VISRES.2016.09.010

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-Specific Delay Activity in Human Primary Visual Cortex. *Psychological Science*, *20*(2), 207–214. http://doi.org/10.1111/j.1467-9280.2009.02276.x

Silver, M. A., Ress, D., & Heeger, D. J. (2005). Topographic Maps of Visual Spatial Attention in Human Parietal Cortex. *Journal of Neurophysiology*, *94*(2), 1358–1371. http://doi.org/10.1152/jn.01316.2004

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219. http://doi.org/10.1016/J.NEUROIMAGE.2004.07.051

Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring Latent Visual Working Memory Representations in Human Cortex. *Neuron*, *91*(3), 694–707. http://doi.org/10.1016/j.neuron.2016.07.006

Sprague, T. C., Itthipuripat, S., Vo, V. A., & Serences, J. T. (2018). Dissociable signatures of visual salience and behavioral relevance across attentional priority maps in human cortex. *Journal of Neurophysiology*, *119*(6), 2153–2165. http://doi.org/10.1152/jn.00059.2018

Sprague, T. C., & Serences, J. T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*, *16*, 1879. Retrieved from https://doi.org/10.1038/nn.3574

Sreenivasan, K. K., Vytlacil, J., & D'Esposito, M. (2014). Distributed and Dynamic Storage of Working Memory Stimulus Information in Extrastriate Cortex. *Journal of Cognitive Neuroscience*, *26*(5),

1141–1153. http://doi.org/10.1162/jocn_a_00556

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405. http://doi.org/10.1016/j.tics.2015.05.004

Sugase-Miyamoto, Y., Liu, Z., Wiener, M. C., Optican, L. M., & Richmond, B. J. (2008). Short-term memory trace in rapidly adapting synapses of inferior temporal cortex. *PLoS Computational Biology*, *4*(5). http://doi.org/10.1371/journal.pcbi.1000073

van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. L. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *ELife*, *7*, 1–25. http://doi.org/10.7554/eLife.38677

Vernet, M., Quentin, R., Chanes, L., Mitsumasu, A., & Valero-Cabré, A. (2014). Frontal eye field, where art thou? Anatomy, function, and non-invasive manipulation of frontal regions involved in eye movements and associated cognitive operations . *Frontiers in Integrative Neuroscience* . Retrieved from https://www.frontiersin.org/article/10.3389/fnint.2014.00066

Wang, L., Mruczek, R. E. B., Arcaro, M. J., & Kastner, S. (2015). Probabilistic maps of visual topography in human cortex. *Cerebral Cortex*, *25*(10), 3911–3931. http://doi.org/10.1093/cercor/bhu277

Warden, M. R., & Miller, E. K. (2007). The representation of multiple objects in prefrontal neuronal delay activity. *Cerebral Cortex*, *17*(SUPPL. 1). http://doi.org/10.1093/cercor/bhm070

Warden, M. R., & Miller, E. K. (2010). Task-Dependent Changes in Short-Term Memory in the Prefrontal Cortex. *Journal of Neuroscience*, *30*(47), 15801–15810. http://doi.org/10.1523/jneurosci.1569-10.2010

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871. http://doi.org/10.1038/nn.4546

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *NeuroImage*, *14*(6), 1370–1386. http://doi.org/10.1006/NIMG.2001.0931

Xu, Y. (2007). The Role of the Superior Intraparietal Sulcus in Supporting Visual Short-Term Memory for Multifeature Objects. *The Journal of Neuroscience*, *27*(43), 11676 LP-11686.

http://doi.org/10.1523/JNEUROSCI.3545-07.2007

Xu, Y. (2018). The posterior parietal cortex in adaptive visual processing. *Trends in neurosciences*, *41*(11), 806-822.

Xu, Y. (2020). Revisit once more the sensory storage account of visual working memory. *Visual Cognition*

Xu, Y. & Chun, M.M. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* **440**, 91–95 (2006).

Yu, Q., Teng, C., & Postle, B. R. (2020). Different states of priority recruit different neural representations in visual working memory. *PLoS Biology*, *18*(6), e3000769.

Yu, Q., & Shim, W. M. (2017). Occipital, parietal, and frontal cortices selectively maintain task-relevant features of multi-feature objects in visual working memory. *NeuroImage*, *157*, 97–107. http://doi.org/10.1016/J.NEUROIMAGE.2017.05.055

Yu, Q., & Shim, W. M. (2019). Temporal-order-based attentional priority modulates mnemonic representations in parietal and frontal cortices. *Cerebral Cortex*, *29*(7), 3182-3192

Zanto, T. P., Rubens, M. T., Thangavel, A., & Gazzaley, A. (2011). Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nature Neuroscience*, *14*, 656. Retrieved from https://doi.org/10.1038/nn.2773

Supplementary material

Supplementary Table 1. Delay Interval: Comparisons against chance. All t-tests were two-tailed.

	Within-relevance Category Decoding (Chance: 33.33%)				
		Mean	SE	T	p	Cohen's D
VC	Current	40.49	1.44	5.1	0.000	1.02
	Prospective	41.90	1.31	6.6	0.000	1.33
	irrelevant	37.65	1.48	2.7	0.012	0.54
pFs	Current	43.99	0.99	11.3	0.000	2.25
	Prospective	41.46	1.38	6.1	0.000	1.21
	irrelevant	39.14	1.03	5.6	0.000	1.12
IPS	Current	39.55	1.02	6.0	0.000	1.20
	Prospective	38.39	1.44	3.5	0.002	0.69
	irrelevant	34.69	1.52	0.7	0.511	0.13
FEF	Current	35.77	1.35	1.7	0.107	0.34
	Prospective	33.58	1.20	-0.2	0.987	-0.004
	irrelevant	35.01	1.04	1.5	0.156	0.29
lPFC	Current	34.86	1.12	1.2	0.256	0.23
	Prospective	33.67	1.48	-0.04	0.969	-0.01
	irrelevant	32.91	1.13	-0.6	0.572	-0.12
	Cross-relevance Category Decoding (Chance: 33.33%)					
		Mean	SE	t	p	Cohen's D
VC	Current-Prospective	41.88	1.42	6.3	0.000	1.26
	Current-Irrelevant	41.12	1.46	5.4	0.000	1.08
	Prospective-Irrelevant	38.51	1.16	4.4	0.000	0.88
pFs	Current-Prospective	43.32	1.01	10.3	0.000	2.06
	Current-Irrelevant	41.85	1.05	8.2	0.000	1.64
	Prospective-Irrelevant	40.73	1.07	7.1	0.000	1.42
IPS	Current-Prospective	39.45	1.08	5.8	0.000	1.16
	Current-Irrelevant	39.06	1.15	5.0	0.000	1.00
	Prospective-Irrelevant	37.46	0.99	4.1	0.000	0.82
FEF	Current-Prospective	35.30	0.98	1.8	0.081	0.36
	Current-Irrelevant	35.51	0.68	3.1	0.005	0.62
	Prospective-Irrelevant	33.97	0.82	0.6	0.535	0.12
lPFC	Current-Prospective	36.45	0.75	4.0	0.000	0.80
	Current-Irrelevant	34.29	0.81	1.0	0.320	0.20
	Prospective-Irrelevant	34.45	1.28	0.6	0.526	0.13
	Status decoding (Chance: 50%)					

		Mean	SE	*t*	*p*	Cohen's D
VC	Current-Prospective	51.79	1.07	1.5	0.141	0.30
	Current-Irrelevant	52.96	0.94	3.1	0.005	0.62
	Prospective-Irrelevant	53.98	1.05	3.6	0.001	0.73
pFs	Current-Prospective	52.85	1.10	2.5	0.019	0.50
	Current-Irrelevant	58.33	0.90	9.5	0.000	1.90
	Prospective-Irrelevant	54.92	0.75	6.6	0.000	1.32
IPS	Current-Prospective	56.87	0.91	7.6	0.000	1.53
	Current-Irrelevant	60.61	1.26	8.6	0.000	1.71
	Prospective-Irrelevant	58.75	0.89	10.2	0.000	2.04
FEF	Current-Prospective	54.59	0.93	4.9	0.000	0.99
	Current-Irrelevant	59.72	1.16	8.6	0.000	1.72
	Prospective-Irrelevant	55.77	1.14	5.0	0.000	1.00
lPFC	Current-Prospective	56.55	1.07	6.2	0.000	1.24
	Current-Irrelevant	61.98	1.35	9.2	0.000	1.84
	Prospective-Irrelevant	59.44	1.04	9.3	0.000	1.87

Supplementary Table 2. Search 1 Interval: Comparisons against chance. All t-tests were two-tailed.

		Mean	SE	t	p	Cohens' D
Within relevance Category Decoding (Chance: 33.33%)
VC	Current	40.64	1.47	4.8	0.000	0.95
	Prospective	35.63	1.18	1.8	0.093	0.35
	Irrelevant	33.16	1.31	-0.4	0.715	-0.08
pFs	Current	49.51	1.71	9.9	0.000	1.98
	Prospective	37.12	1.13	3.3	0.003	0.66
	Irrelevant	33.91	0.43	0.6	0.671	0.09
IPS	Current	41.73	1.14	7.4	0.000	1.55
	Prospective	35.19	1.47	1.3	0.109	0.21
	Irrelevant	33.46	1.08	0.1	0.454	0.02
FEF	Current	37.16	1.32	2.7	0.014	0.53
	Prospective	33.38	1.13	-0.2	0.869	-0.03
	Irrelevant	29.68	1.04	-3.6	0.001	-0.72
lPFC	Current	37.88	1.49	2.8	0.009	0.57
	Prospective	33.51	1.17	-0.1	0.954	-0.01
	Irrelevant	33.04	0.96	-0.5	0.636	-0.10

		Mean	SE	t	p	Cohens' D
Cross-relevance Category Decoding (Chance: 33.33%)
VC	Current-Prospective	31.98	0.80	-1.8	0.080	-0.37
	Current-Irrelevant	32.64	0.83	-1.0	0.343	-0.20
	Prospective-Irrelevant	35.28	0.86	2.2	0.041	0.43
pFs	Current-Prospective	28.84	0.76	-5.8	0.000	-1.17
	Current-Irrelevant	31.79	0.82	-2.0	0.057	-0.40
	Prospective-Irrelevant	34.98	0.83	1.9	0.070	0.38
IPS	Current-Prospective	32.40	0.82	-1.3	0.213	-0.26
	Current-Irrelevant	32.07	0.82	-1.7	0.107	-0.34
	Prospective-Irrelevant	35.00	1.17	1.2	0.237	0.24
FEF	Current-Prospective	31.93	0.72	-2.1	0.049	-0.42
	Current-Irrelevant	33.07	0.63	-0.5	0.607	-0.11
	Prospective-Irrelevant	34.19	0.88	0.8	0.419	0.16
lPFC	Current-Prospective	31.70	0.84	-2.1	0.050	-0.42
	Current-Irrelevant	32.75	0.84	-0.8	0.409	-0.17
	Prospective-Irrelevant	33.46	0.98	-0.05	0.963	-0.01

\multicolumn{7}{c}{Status decoding (Chance: 50%)}

		Mean	SE	t	p	Cohens' D
VC	Current vs. Prospective	64.53	1.28	11.9	0.000	2.37
	Current vs. Irrelevant	66.22	0.93	18.5	0.000	3.70
	Prospective vs. Irrelevant	54.46	1.04	4.2	0.000	0.83
pFs	Current vs. Prospective	64.94	1.16	13.4	0.000	2.68
	Current vs. Irrelevant	62.98	1.10	12.4	0.000	2.49
	Prospective vs. Irrelevant	54.25	1.03	4.1	0.000	0.81
IPS	Current vs. Prospective	66.79	1.20	14.8	0.000	2.97
	Current vs. Irrelevant	65.61	1.13	14.5	0.000	2.90
	Prospective vs. Irrelevant	54.88	1.15	4.2	0.000	0.84
FEF	Current vs. Prospective	58.62	1.27	7.0	0.000	1.39
	Current vs. Irrelevant	57.37	0.89	8.4	0.000	1.69
	Prospective vs. Irrelevant	55.44	1.06	5.2	0.000	1.04
lPFC	Current vs. Prospective	63.88	1.36	10.7	0.000	2.14
	Current vs. Irrelevant	60.74	0.85	13.1	0.000	2.62
	Prospective vs. Irrelevant	56.46	0.91	7.2	0.000	1.44

Supplementary Table 3. Search 2 Interval: Comparisons against chance. All t-tests were two-tailed.

		Mean	SE	t	p	Cohens' D
Within relevance Category Decoding (Chance: 33.33%)						
VC	Current	34.57	1.60	0.5	0.608	0.10
	Prospective	43.01	1.41	7.1	0.000	1.42
pFs	Current	34.48	1.00	1.1	0.132	0.20
	Prospective	49.05	1.25	12.5	0.000	2.69
IPS	Current	34.37	1.14	0.7	0.496	0.14
	Prospective	39.70	1.10	5.7	0.000	1.14
FEF	Current	34.94	1.08	1.3	0.199	0.26
	Prospective	33.28	1.31	-0.3	0.780	-0.06
lPFC	Current	32.37	0.97	-1.2	0.252	-0.24
	Prospective	33.83	1.21	0.2	0.853	0.04

Cross-relevance Category Decoding (Chance: 33.33%): Current-Prospective

	Mean	SE	t	p	Cohens' D
VC	29.75	0.83	-4.4	0.000	-0.88
pFs	27.78	0.84	-6.5	0.000	-1.31
IPS	30.91	0.92	-2.8	0.011	-0.54
FEF	32.80	0.86	-0.8	0.449	-0.16
lPFC	33.63	0.63	0.4	0.723	0.07

Status decoding (Chance: 50%)

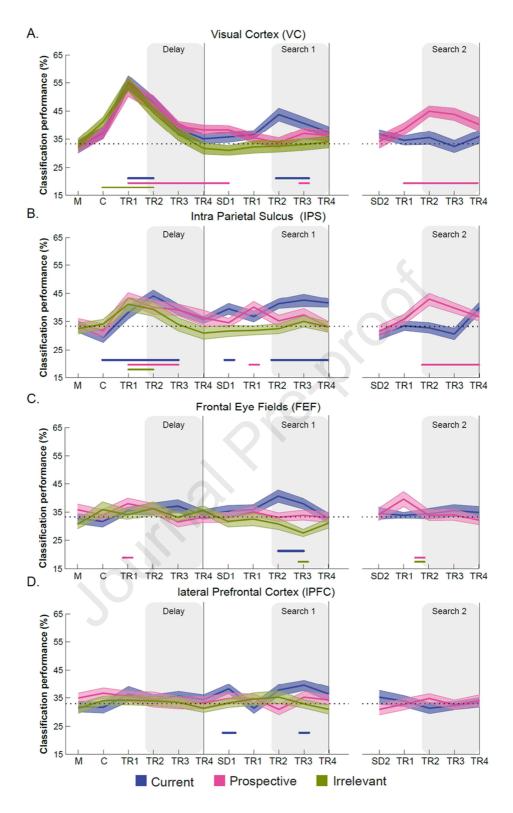|	Mean	SE	t	p	Cohens' D
VC	65.53	1.08	15.0	0.000	3.01
pFs	63.29	1.16	12.0	0.000	2.40
IPS	68.33	1.19	16.4	0.000	3.29
FEF	57.72	1.25	6.2	0.000	1.24
lPFC	65.57	1.17	14.3	0.000	2.87

Figure S1.Time course of Object category decoding (Within-relevance) for VC, IPS, FEF and lPFC. Shaded blue and pink areas indicate standard error. Blue, pink and green horizontal lines at the bottom of the graphs indicate time points that significantly differ from chance in the one-tailed one sample t-test (p < 0.05, uncorrected).
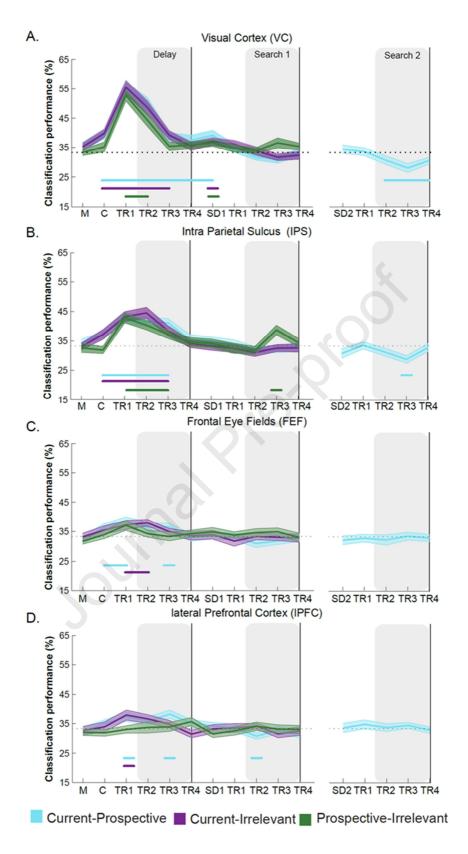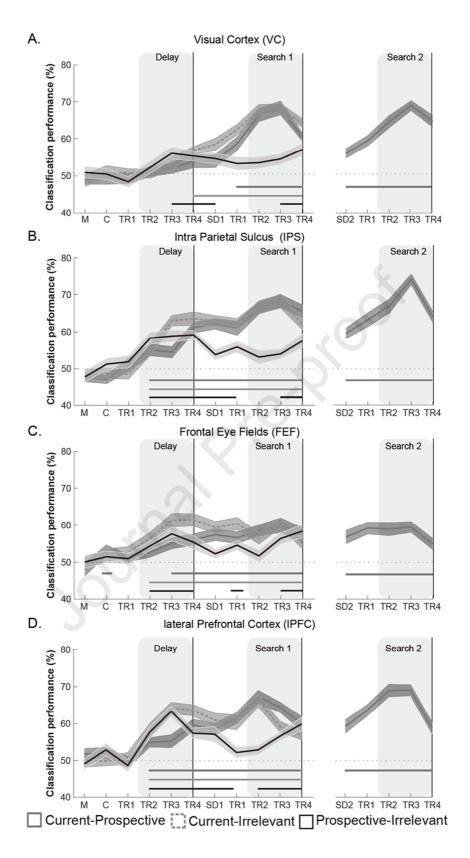
Figure S2.Time course of Object category decoding (Cross-relevance) for VC, IPS, FEF and lPFC. Shaded blue and pink areas indicate standard error. Blue, pink and green horizontal lines at the bottom of the graphs indicate time points that significantly differ from chance in the two-tailed one sample t-test (p < 0.05, uncorrected).
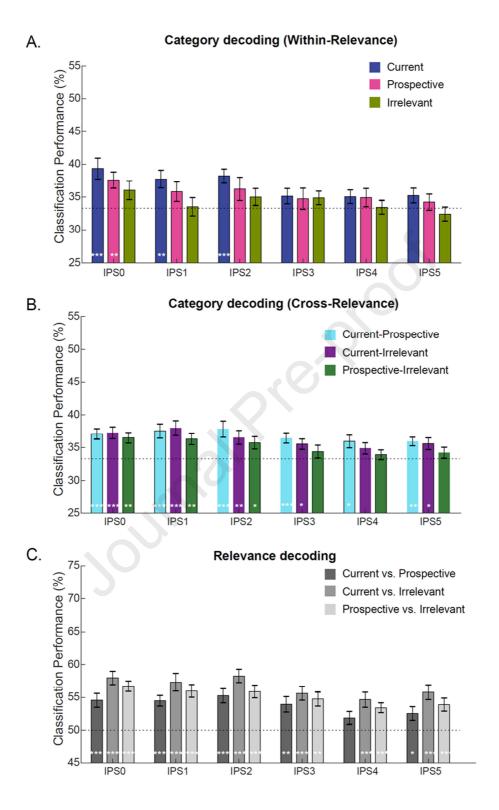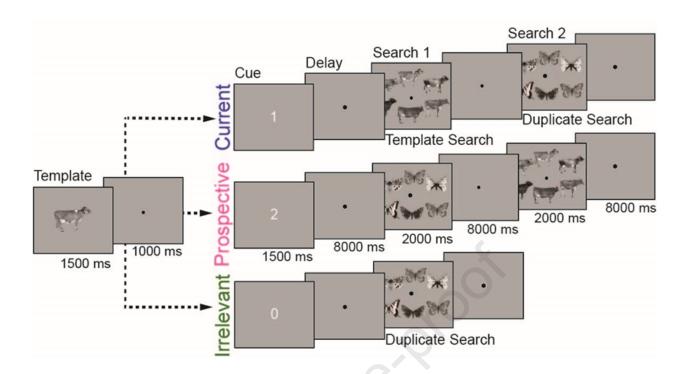
Figure S3.Time course of Status decoding for VC, IPS, FEF and lPFC. Data from pFs were also included for comparison purposes. Shaded blue and pink areas indicate standard error. Blue, pink and green horizontal lines at the bottom of the graphs indicate time points that significantly differ from chance in the one-tailed one sample t-test (p < 0.05, uncorrected).
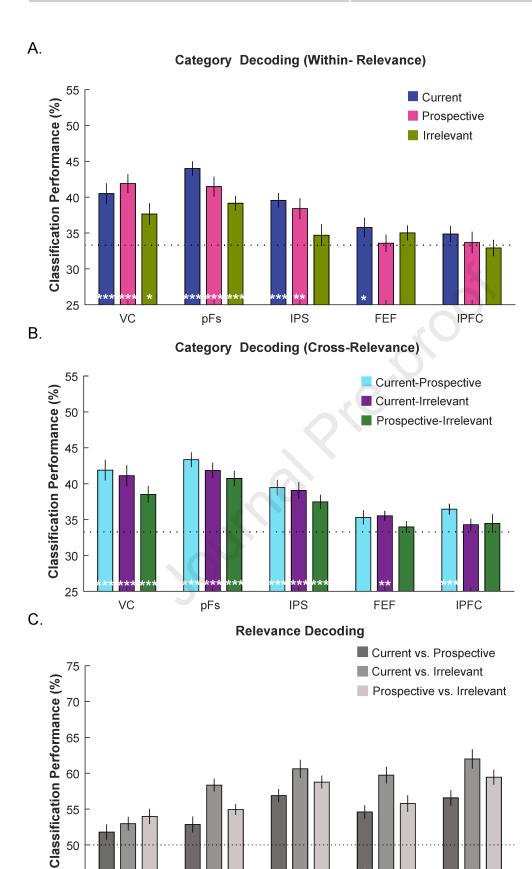
Figure S4. Delay interval: Object category decoding (Within-relevance and cross-relevance) and status decoding for different IPS subsections (IPS0 to IPS5). (A) Category classification performance within each relevance condition (Current, Prospective Irrelevant). (B) Category classification performance using cross-relevance decoding. (C) Status decoding. Error bars indicate Standard Error (SE) of the mean, N= 25. * p<0.05, **p < .01, ***p< .001.

48

A.

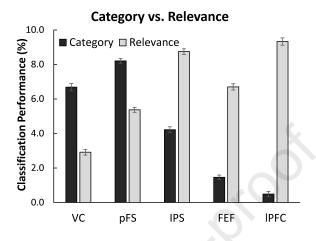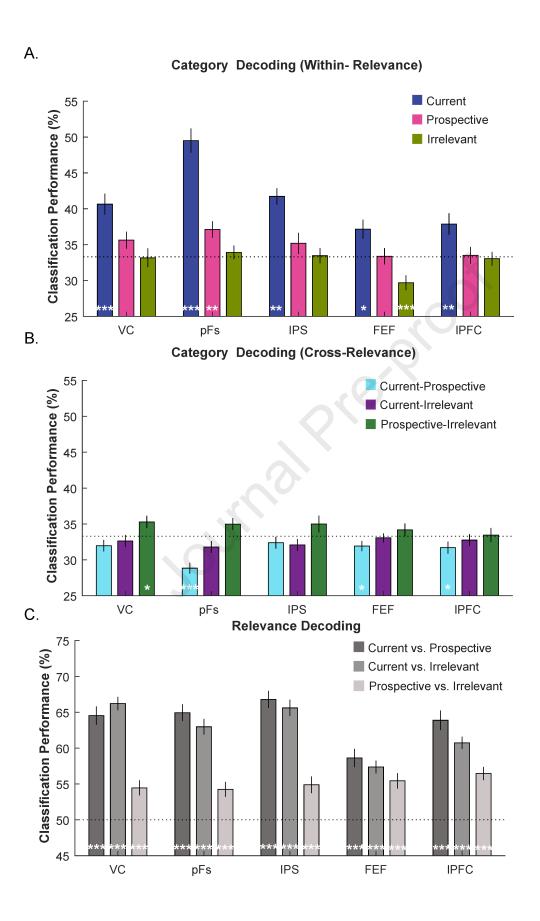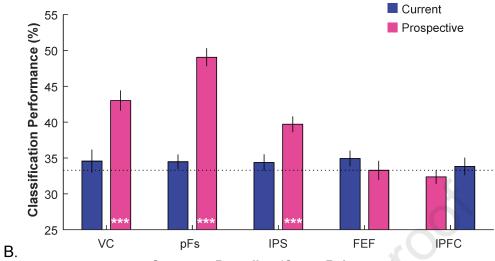**Category Decoding (Within- Relevance)**

B.

Category Decoding (Cross-Relevance)

C.

Relevance Decoding

Category vs. Relevance

A.

Category Decoding (Within- Relevance)

B.

Category Decoding (Cross-Relevance)

C.

Relevance Decoding

A.

Category Decoding (Within- Relevance)

B.

Category Decoding (Cross-Relevance)