Visual Cognition

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pvis20

No difference in prior representations of what to attend and what to ignore

Michel Failing, Gilles de Hollander, Stefan Pollmann & Christian N. L. Olivers

To cite this article: Michel Failing, Gilles de Hollander, Stefan Pollmann & Christian N. L. Olivers (23 Aug 2024): No difference in prior representations of what to attend and what to ignore, Visual Cognition, DOI: 10.1080/13506285.2024.2338605

To link to this article: https://doi.org/10.1080/13506285.2024.2338605

VISUAL COGNITION
https://doi.org/10.1080/13506285.2024.2338605

Routledge
Taylor & Francis Group

No difference in prior representations of what to attend and what to ignore

Michel Failing[a,b], Gilles de Hollander[c,d], Stefan Pollmann[e,f] and Christian N. L. Olivers[g,h]

[a]Charité – Universitätsmedizin Berlin, Berlin, Germany; [b]Berlin Institute of Health, Berlin, Germany; [c]Zürich Center for Neuroeconomics, Department of Economics, University of Zürich, Zürich, Switzerland; [d]Spinoza Centre for Neuroimaging, Royal Academy of Sciences, Amsterdam, The Netherlands; [e]Department of Experimental Psychology, Institute of Psychology, Otto-Von-Guericke University, Magdeburg, Germany; [f]Center for Behavioral Brain Sciences, Magdeburg, Germany; [g]Department of Experimental and Applied Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; [h]Institute of Brain and Behavior Amsterdam, Amsterdam, The Netherlands

ABSTRACT
Adaptive behaviour requires the selection of relevant, and the ignoring of irrelevant visual information. Previous work has shown that prior information about target (i.e., relevant) and non-target (i.e., irrelevant) objects facilitates such selection, presumably by enabling observers to create a working memory template for or against such objects. Using a cued visual search task, here we aimed to investigate whether, and how, target and non-target templates differ in terms of neural representation, as measured through fMRI activity patterns. Two hypotheses were tested: (1) while target representations are activated, non-target representations are suppressed prior to search, and thus these representations differ, and (2) target and non-target templates share similar initial representations but involve different control signals in anticipation of, or upon encountering the search display. Behaviourally, both target and non-target cues facilitated search compared to baseline. Using multivariate pattern and representational similarity analyses, the findings revealed little support for the suppression hypothesis. Posterior brain regions primarily coded for the category content of the template and did so similarly for target and non-target information. Frontal cortical areas, particularly the lateral prefrontal cortex and frontal eye fields, instead showed sensitivity to the status of the template, regardless of the object category it belonged to. These results are most consistent with the hypothesis that prior to selection, target and non-target information are represented similarly in terms of content, but differ in terms of the associated control signals.

ARTICLE HISTORY
Received 1 June 2023
Accepted 28 March 2024

KEYWORDS
attention; visual search; distractor suppression; visual working memory; fMRI; cognitive control

A fundamental obstacle in being able to interact with the environment in a meaningful way is the sheer amount of complex visual stimulation that is constantly hitting the retina. For adaptive behaviour, mechanisms of attention selectively prioritize certain information for more in-depth processing above and beyond other information. Such selectivity becomes evident in for instance visual search tasks, in which observers look for a relevant (target) object while trying to ignore irrelevant (non-target) objects. For example, when looking for the car key fob on the kitchen counter, one might prioritize small black objects over other, irrelevant objects such as a coffee mug, the note your partner left for you, and a sandwich leftover.

A large body of research has shown that biasing selection towards target features occurs already in advance, prior to the onset of to-be-searched visual information (e.g., Battistoni et al., 2017; Folk et al., 1992; Olmos-Solis et al., 2018; van Loon et al., 2017). Indeed, in line with some of the most influential theories of attention (e.g., Bundesen, 1990; Desimone & Duncan, 1995; Duncan & Humphreys, 1989), representations of target features have been found in individual as well as across populations of neurons already prior to presentation of the target stimulus (e.g., Chawla et al., 1999; Chelazzi et al., 1993, 1998; Giesbrecht et al., 2006; Peelen & Kastner, 2011; Stokes et al., 2009). By their functional role, such *target templates*, as they have been referred to, must be flexible in

CONTACT Christian N. L. Olivers c.n.l.olivers@vu.nl Faculty of Behvioural and Movement Sciences, Vrije Universiteit Amsteram, Van der Boechorststraat 7, 1081 BT, Amsterdam, The Netherlands

Supplemental data for this article can be accessed online at https://doi.org/10.1080/13506285.2024.2338605

nature, in that they are specifically tuned to the sought-for features at hand, and maintained until no longer necessary (Gayet et al., 2017; Grubert et al., 2017; Olivers et al., 2011; Olivers & Eimer, 2011). Mechanistically, such target templates have been hypothesized to pre-activate or pre-sensitize the very same neurons that code for the target features when presented as a stimulus, thus facilitating its processing and prioritizing it for selection (e.g., Desimone & Duncan, 1995).

While the existence of template representations that proactively and flexibly enhance target objects (or target-related features) is widely accepted, there is an ongoing debate about whether such templates can also be set up to proactively, and flexibly, ignore irrelevant objects or their features (see also Chelazzi et al., 2019, for a recent review). Such "rejection templates" or "non-target templates" (the latter of which will be used from here on) would in principle be beneficial to performance when there is useful information about non-target features prior to search.[1] Evidence from behavioural studies indicates that observers can indeed effectively and flexibly use prior information on what feature to ignore in a subsequent search display, as search becomes more efficient with than without such information (e.g., Carlisle & Nitka, 2019; Arita et al., 2012; Conci et al., 2019; Woodman & Luck, 2007; Zhang et al., 2020) – although benefits tend to be smaller than when target information is provided.

A crucial question is then how such efficiency is achieved from non-target templates. A first hypothesis states that providing non-target information results in selective *suppression* of its neural representation already in advance (cf. Arita et al., 2012; Woodman & Luck, 2007). Depending on the strength of the suppression, non-target templates may become either attenuated compared to target templates (cf. Reeder et al., 2018), silent (cf. Olivers et al., 2011), or even be suppressed below baseline (cf. van Loon et al., 2018; Wan et al., 2020; Yu et al., 2020). Under these possibilities, corresponding patterns of brain activity for non-target templates would then be expected to respectively show a weaker correlation, zero correlation, or an anti-correlation with target templates. In all cases, the content representation of the non-target template differs in terms of activity patterns from that for the target template. We will collectively refer to these possibilities as the *prior suppression hypothesis*. The main alternative hypothesis is that non-target templates are, in terms of both content and strength, represented in exactly the same way as regular target templates, but come with different *control* operations attached. Such a control operation could be a distractor-specific plan or intention to *re*actively deal with a distractor once it is actually encountered in the search display (cf. Gaspelin & Luck, 2018; Moher & Egeth, 2012; Sawaki & Luck, 2013). It could also represent an overall signal gating the entire perceptual input at search onset, or gating the overall speed of responding to that input, in order to prevent too rapid and therefore inaccurate orienting towards distractors, and let top-down processes prevail. For example, in their EEG study, de Vries et al. (2019) found little difference in how target- versus distractor-specific representations were prioritized before the search display, as reflected in lateralized posterior alpha power. Instead, subsequent eye movement activity during the search itself showed that participants first attended to the distractor, before turning away, in line with what Moher and Egeth (2012) proposed. However, compared to target information, providing distractor information did result in modulations control-related signals (overall posterior alpha as well as mid-frontal theta signals), signals that were predictive of how strongly both targets and distractors in the display then captured attention. De Vries et al. therefore hypothesized that this activity represented central control signals which generally gated the perceptual input at search display onset, thus preventing attentional capture in general, rather than in a distractor-specific manner (see also Noonan et al., 2018; Rajsic et al., 2020; van Zoest et al., 2021 for similar findings and accounts). Finally, control operations could represent a plan to *recode* the non-target information into target information once the display has appeared. For example, Beck and Hollingworth (2015) as well as Becker et al. (2015) argued that prior information on what will be irrelevant can be used to derive what *is* relevant as soon as the search display appears, allowing for a recoding of attentional priorities without the need for an inhibitory template. In the present study, we will not be able to discriminate between these versions, and we will collectively refer to them as the *control hypothesis*. For all cases, this hypothesis predicts that the template's content representation is not affected as such, and thus this content representation should not differ for target and non-target templates.

Instead, such control signals would be expected to indicate only the *status* of the information – that is, is this going to be a prospective target or a non-target – and such signals would therefore be expected to emerge in domain-general executive control networks involving frontal and/or parietal cortex.

Two earlier studies used fMRI to compare target and non-target templates prior to the search (Reeder et al., 2018; Reeder et al., 2017). In these studies, participants were required to identify the orientation of a target letter appearing among multiple pseudo-letters. Half of these stimuli appeared inside disks of a particular colour, while the other half appeared in disks of another colour. Crucially, participants were given a cue before search onset which either indicated the colour in which the target letter would appear (target cue), the colour in which it would not appear (non-target cue) or a colour that would not be present in the search display (neutral cue). At the behavioural level, Reeder and colleagues observed that both target and non-target templates benefitted search, with target templates being the most effective. Univariate analysis of the neural activity leading up to the search showed an overall lower BOLD signal in response to non-target cues relative to the neutral cue baseline in early visual areas (V1–V3), while target cues showed a larger BOLD response relative to neutral. Additionally, and probably as a direct consequence of the overall reduced BOLD signal, a representational similarity analysis (Kriegeskorte et al., 2008) of the colour representations in the early visual cortex revealed that target templates were represented more distinctly from neutral feature representations, while no such difference was found when comparing non-target and neutral representations. Such less pronounced activity for nontargets would be consistent with prior suppression. However, given that non-target cues are generally less effective than target cues, the reduced overall activity may also be due to a reduced number of trials on which participants are willing to activate a representation that they may experience as less effective for the task. The case for truly inhibitory non-target templates would be stronger if we were to find patterns of activity that are suppressed below the baseline and that would thus anti-correlate with target templates. This is what we set out to test in the current study.

Here we re-addressed the question of how the functional advantage of using non-target templates in visual search manifests itself on a neural level, using a multivariate analysis approach in combination with a stimulus set consisting of different object categories, and for which previous studies had observed goal-dependent changes in representation (Harel et al., 2014; Lee et al., 2013; van Loon et al., 2018). To differentiate between the prior suppression and control hypotheses, we asked participants to perform a difficult search task for a T-shaped symbol among L-shaped symbols, which were presented on top of an array of objects drawn from one of three different object categories (cow, dresser or skate), while we measured brain activity using fMRI. We chose object categories over colours (as used by Reeder et al., 2018; Reeder et al., 2017), because multivariate decoding of object categories tends to be more reliable than of decoding colour categories, therefore increasing the chances of uncovering potentially subtle differences (see also Seidl et al., 2012). Moreover, other work has shown that changes in mnemonic representational patterns can be found for these stimuli depending on the task goal (Harel et al., 2014; Lee et al., 2013; van Loon et al., 2018). The procedure is illustrated in Figure 1 (A). Importantly, at the beginning of each trial the participants received information in the form of a template object meant to guide their search strategy. For this purpose, the template object was followed by either one of four different cues: a *target*, *non-target*, *neutral* and *drop* cue. In the *target cue* condition, the search target would always appear on an instance of a memorized object. In the *non-target cue* condition, participants were informed that while instances of the memorized object would be present in the search display, the target would certainly not appear on any of them. In the *neutral cue* condition, participants were informed that instances of the memorized object would be present in the search display but not whether they would contain the target (i.e., be relevant) or only distractors (i.e., be irrelevant). This information was only relayed at search onset, through a second cue, which could be target or non-target as per above. In other words, in this condition, it was still useful for participants to memorize the object, but during the delay, it was not yet clear how to use it. Finally, in the *drop cue* condition, participants were informed that the

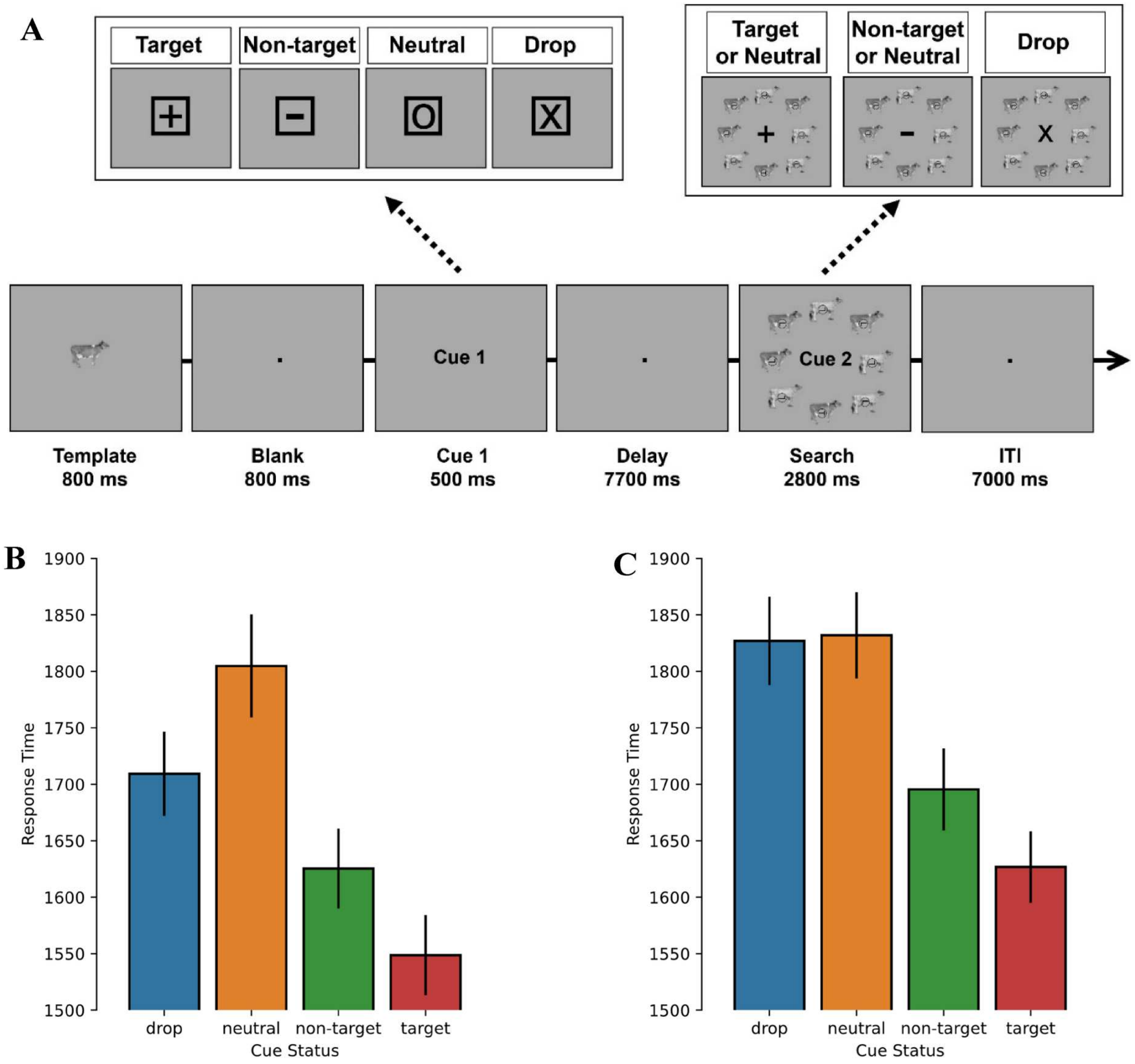

Figure 1. (A) Trial sequences of the experiment. On each trial, participants performed a search task for a T-shaped symbol among L-shaped symbols presented on top of an array of greyscale objects. These objects were drawn from one of three different object categories (cow, dresser or skate). Prior to the search display onset, participants received information in the form of a template object meant to guide their search strategy. The content of the information was conveyed by means of a cue presented shortly after the template object and was systematically manipulated creating four different cue conditions: in the *target cue* condition, the search target would always appear on the memorized object. In the *non-target cue* condition, participants were informed that while the memorized object would be present in the search display the target would certainly not appear on it. In the *neutral cue* condition, participants were informed that the memorized object would be present in the search display but that only a second cue, appearing at search onset, would indicate whether the target would appear inside the memorized object (i.e., target cue) or not (i.e., non-target cue). Finally, in the *drop cue* condition, participants were informed that the memorized object would not appear in the search display and thus could be dropped from memory. (B) Behavioral results of the behavioral session. Average response time for each cue condition. (C) Behavioural results of the fMRI session. Average response time for each cue condition. Error bars here and in all other figures indicate standard error of the mean.

memorized object would not appear in the search display and thus could be dropped from memory altogether. The latter two conditions served as baseline conditions relative to which we exploratively compared the representational status of target and non-target templates, to see if non-target templates actually involve below-baseline activity (MacLeod, 2007). However, the main comparison of interest was the one between target and non-target templates. Multivariate pattern activity was then taken from early visual areas (V1-V3), lateral occipital cortex (LOC), posterior fusiform area (pF), intraparietal sulcus (IPS), frontal eye fields (FEF) and lateral prefrontal cortex (lPFC) to assess where and how the content of the template (i.e., its object category membership), but also the status of the template (i.e., whether it is a target, non-target, neutral or dropped template) is represented. The choice of the different ROIs was motivated by previous studies which indicated these areas as functionally important to the maintenance of visual information as well as task-specific modulations thereof, including planning of behaviour and dealing with distractor interference (Baeck et al., 2013; Barbey et al., 2013; Bettencourt & Xu, 2016; Christophel et al., 2018; Cosman et al., 2018; Grill-Spector et al., 2001; Lee et al., 2013; van Loon et al., 2018; Lorenc et al., 2018; Mansouri et al., 2017; Marini et al., 2016; Olmos-Solis et al., 2021; Rademaker et al., 2019).

There were two main analyses of interest. First, *category* decoding analyses focused on the content of the template representation, and whether such representation differed for target versus non-target templates. If nontargets are represented through inhibitory templates we should observe at least much weaker decoding and reduced similarity for non-target templates compared to target templates, but could even observe reversals of representation if object-selective voxels are suppressed below their usual baseline. Second, *status* decoding ignores the object category, but provides a measure of how sensitive a brain region is to the status of the template *per se* – that is, whether it is a target or a non-target. Because this decoding analysis by definition generalizes across categories, it is more likely to reflect general control signals, rather than stimulus-specific modulations.

Materials and methods

Participants

Data were initially obtained from twenty-four volunteers. The sample size was predetermined on the basis of van Loon et al. (2018). Data from two participants were excluded from the analysis. For one of these participants, data collection was aborted prematurely due to the experience of a high degree of discomfort in the MRI scanner. Another participant was excluded due to excessive movement throughout the entire scanning session (framewise displacement: maximum regularly around 23 and run averages around 6). This left a sample of 22 participants (13 female, M = 24.8 ± SD = 4.4 years old). All participants had reported normal or corrected-to-normal vision and were naïve as to the purpose of this study. All participants gave written consent prior to the start of the experiment. The study was approved by the Scientific and Ethics Review Board of the Faculty of Behavioural and Movement Sciences of the Vrije Universiteit Amsterdam (The Netherlands).

Stimuli

Stimulus presentation was controlled with OpenSesame (Mathôt et al., 2012) and stimuli were always presented against a uniform grey background (RGB-values: 170, 170, 170). Participants were shown real-world objects in the form of greyscale photographs depicting one of four exemplars from one of three object categories (cow, dresser, or roller skate; approximately 2.43° by 2.43° visual angle). These object categories were selected to have maximal dissimilarity in representational space (see Harel et al., 2014). In the search display, each of the object pictures had a circular recess in its centre (black outline with 0.2° radius) that contained either a T- or L-shaped symbol. The T-shaped symbol only appeared exactly once in the search display and was rotated either 90° clockwise or counterclockwise from vertical, while the remaining objects contained L-shaped symbols in one of four rotations (0°, 90°, 180°, and 270°). All letter-shaped symbols were made up of two equally long line segments (0.21°). All pictures in the search display were presented on an imaginary circle at an equal distance (3° radius centre-to-centre). The locations of all eight pictures and their T- or L-shaped symbol were randomized. Depending on the condition, the cue display contained either a “+”, “-“, “O” or “X” symbol, respectively indicating target, non-target, neutral, and drop. Since these symbolic cues were difficult to match in pixel size without significantly compromising their overall comparability in other aspects (such as size), each cue was shown inside a rectangle (1.33° by 1.33°) to minimize the relative influence of the difference in pixels on the BOLD signal. The very same cues also appeared in the centre of the search display. In between displays, a black fixation dot was visible at the centre of the display.

Procedure and design

Figure 1 shows the trial sequence. Each trial contained a template, cue and search display. A trial started with the 800 ms long presentation of the to-be-memorized template followed by a fixation screen for another 800 ms. This template was a picture of an exemplar taken from one of the object categories (cow, dresser, skate). Participants were instructed to memorize the template as veridically as possible and to avoid using semantic or verbal labels. To promote this, it was insufficient for participants to simply memorize the object category, because all objects in the search display would be of the same category. Instead, a specific exemplar had to be remembered for the template to be effective. Then, following the presentation of the template, a cue was shown for 500 ms which informed participants about how the template should be used in the upcoming search. The presentation of the cue was followed by a delay period of 7700 ms in which only a central fixation dot was visible. Finally, the search display came on for 2800 ms and was followed by an intertrial interval of another 7000 ms. The search display always contained eight search elements all of which were pictures of one object category (cow, dresser or skate) and each contained a rotated L-shaped object in the centre, except for the target, which contained a rotated T. There were always four pictures of one exemplar and four pictures of another exemplar of the designated object category. Participants were required to indicate the rotation of the T-shaped target by pressing one of two buttons using separate hands for each of them. The response could be given even after the search display disappeared (up until 3500 ms after search display onset). Response time and accuracy feedback were given at the end of each block, at which point participants could also take a break.

The critical design feature was the type of cue that was presented prior to the delay period. There were four different cue types: a *target*, a *non-target*, a *neutral* and a *drop* cue. A target cue indicated that the target in the search display would appear inside a picture of the shown exemplar, while different pictures would only contain distractors. This effectively allowed participants to restrict search to four of the eight elements in the search display, thus making the search significantly easier. In contrast, a non-target cue indicated that the search target would definitely *not* appear inside a picture of the shown exemplar, but always be inside a picture of another (as of yet unknown) exemplar. The two remaining cue conditions (neutral and drop) functioned as baseline conditions. A neutral cue indicated that the cue instructions would be given at the start of the search, in which case the cue type could only be a target or non-target cue. Finally, a drop cue indicated that not only the exemplar but the entire object category would not be present in the search display. This allowed participants to altogether drop the template from memory entirely as it would, if anything, be detrimental to maintain it. In other words, while the neutral condition functioned as a baseline in which category-specific information (e.g., speckled cow) was maintained without status-specific information (target or non-target), the drop condition functioned as a zero-maintenance baseline in which no category-specific nor status-specific information was required to be maintained.

Each participant performed the same task twice, once in a behavioural session outside the scanner, and once in an fMRI session 1–2 days later. In the behavioural session, participants did eight blocks of 24 trials each. This session allowed participants to become well practiced on the task and thus reduced the influence of fatigue and training effects during the fMRI session. To promote learning, participants received trial-by-trial feedback during the behavioural session by briefly turning the fixation dot white following an incorrect or timed-out response. All of the critical task parameters were kept the same between sessions. That is, during the fMRI session, participants performed another 192 trials with 24 trials in each of the eight blocks. Here trial-based feedback was omitted. In each block of each session, every object category was equally often the object category of the template and every cue condition was tested equally often. The order of conditions was randomized within blocks. Across all trials, each exemplar of each category within each cue condition was tested four times.

Apparatus and functional MRI acquisition

Functional magnetic resonance imaging was realized using a 3 T Philips MRI scanner. Prior to the fMRI session, a high-resolution 3D T1-weighted anatomical scan (repetition time [TR] = 8.196 ms, TE = 3.75 ms, flip angle = 90°, FOV = 240 mm × 256 mm × 250 mm, slice thickness = 1 mm, inversion time 950 ms, voxel size = 1 mm^3) was obtained for every participant.

During the experiment, functional images were recorded using a T2*-weighted single-shot gradient echo-planar images sequence (whole-brain coverage, TR = 700 ms, TE = 30 ms, flip angle = 8°, FOV = 216 mm × 130 mm × 216 mm, slice gap = .27 mm, slice thickness = 2.7 mm, voxel size = 2.7 mm^3, mulitband factor = 4). A Cambridge Research Systems LiveTrack AV Eye-tracker was used to monitor fixation during the task, particularly during the search, for all except two participants in which the eye-tracker malfunctioned.

fMRI data analysis

All analysis scripts can be found on GitHub (https://github.com/MichlF/science_projects/tree/master/NoDifferenceInPriorRepresentationsOfWhatToAttendAndWhatToIgnore).

Preprocessing

fMRI data were preprocessed using FMRIPrep (Esteban et al., 2019) a Nipype (Gorgolewski et al., 2011) based tool. Each T1w volume was corrected for intensity non-uniformity using N4BiasFieldCorrection v2.1.0 (Tustison et al., 2010) and skull-stripped using antsBrainExtraction.sh (using the OASIS template). Brain surfaces were reconstructed using recon-all from FreeSurfer v6.0.0 (Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical grey-matter of Mindboggle (Klein et al., 2017). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009) was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 (Avants et al., 2008), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and grey-matter (GM) was performed on the brain-extracted T1w using fast in FSL v5.0.9 (Zhang et al., 2001).

Functional data were slice time corrected using 3dTshift (Cox, 1996) and motion corrected using mcflirt (FSL v5.0.9; Jenkinson et al., 2002). Distortion correction was performed using an implementation of the PEB/PEPOLAR (phase-encoding based / PE-POLARity) technique. This was followed by co-registration to the corresponding T1w using boundary-based registration with nine degrees of freedom using bbregister (FreeSurfer v6.0.0). Motion correcting transformations, field distortion correcting warp, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation.

Physiological noise regressors were extracted by applying CompCor (Behzadi et al., 2007). Principal components were estimated for anatomical CompCor (aCompCor). A mask to exclude signals with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. For aCompCor, six components were calculated within the intersection of the subcortical mask and the union of CSF and WM masks was calculated in T1w space, after their projection to the native space of each functional run. Frame-wise displacement (Power et al., 2014) was calculated for each functional run using the implementation of Nipype. Many internal operations of FMRIPrep use Nilearn (Abraham et al., 2014), principally within the BOLD-processing work-flow. For more details on the preprocessing pipeline, see https://fmriprep.readthedocs.io/en/1.2.0/workflows.html.

Anatomical regions of interest (ROIs)

Analyses were confined to six anatomically defined brain regions that have previously been found to be involved in attention as well as the maintenance and manipulation of VWM content. These regions are the visual cortex (VC), lateral occipital cortex (LOC), the intraparietal sulcus (IPS), posterior fusiform gyrus (pF), the frontal eye fields (FEF) and the lateral prefrontal cortex (lPFC) which is typically defined as consisting of the ventrolateral prefrontal cortex (vlPFC) and dorsolateral prefrontal cortex (dlPFC). More specifically, VC comprised striate (V1) and extrastriate (V2 and V3) areas which were defined as Brodmann area (BA) 17 through 19. Both LOC and IPS were ROIs taken from the MSDL atlas (Varoquaux et al., 2011). For pF, the temporal occipital fusiform cortex was extracted from the Harvard-Oxford Structural Atlas of the FSL package (similar to Olmos-Solis et al., 2021 and van Loon et al., 2018). FEF were defined as BA 8 (e.g., Paus, 1996). Similarly, regions vlPFC and dlPFC were defined based on their associated BAs. Specifically, vlPFC comprised BA 44, 45 and 47, while dlPFC comprised BA 8, 9, and 46. BAs were extracted from the Talairach atlas using Nilearn (Lancaster et al., 2000) and each ROI mask was resampled into MNI space before being applied on a per-participant level.

Deconvolution

To estimate the average fMRI response time course, we ran a deconvolution analysis using nideconv (de Hollander & Knapen, 2018). For each subject and each trial, a general linear model using its closed-form solution ordinary least squares (OLS) was fitted to the data with every TR (.7 secs each, 28 in total) as a regression variable. The design we employed here used very long intertrial intervals (7 seconds), and the time between subsequent cue periods was 11.4 seconds. Therefore, the overlap in BOLD responses of corresponding within-trial events was minimized, precluding issues of collinearity of the responses (Mumford et al., 2012; Prince et al., 2022). Nuisance regressors included the temporal derivatives, motion-related parameters (three regressors each for translation and rotation), framewise displacement, and six aCompCorr regressors. All regressors were convolved with a Fourier basis function set consisting of an intercept and six sine-cosine pairs of increasing frequency. With this method, we derived vectors containing the *t*-value per voxel in each ROI for every participant, run, experimental condition (target, non-target, neutral, and drop), category exemplar (cow, dresser, and skate; four exemplars of each) and TR in every ROI separately. As such, each of these vectors represents the spatial activity pattern evoked at a given TR for a given trial in a given ROI.

Category and status decoding

To shed light on the question whether the status information of the cue (target, non-target, neutral or drop) affected the neural representations of object categories and/or the status itself was reliably represented in the brain, we further analysed the multi-voxel patterns. We conducted three types of decoding: the first two types concerned content decoding, where we trained and tested the classifier on the three object type categories. In one version, *within-condition* category decoding, we decoded the object category information within each status separately (i.e., target, non-target, neutral, dropped). In other words, classifiers were trained and tested on the same status condition (e.g., cows in the target condition vs. dressers in the target condition). For the *cross-condition* category decoding, we trained the classifier on object categories within one status condition (e.g., non-target), but then tested the classifier on another status condition (e.g., target), and vice versa (e.g., cows in the target condition vs. dressers in the non-target condition). This was done for all six cue condition combinations (drop vs. non-target, drop vs. neutral, drop vs. target, neutral vs. non-target, non-target vs. target and neutral vs. target). This approach allowed us to assess if category information was stored in the same format (predicting above-chance classification), the opposite or anti-correlated format (predicting below-chance classification), or uncorrelated formats (predicting no correlation and thus classification at chance) for each of these combinations. Finally, the third type, *status decoding*, assessed which areas were sensitive to status information per se, regardless of category. Here, the classifier was trained and tested on the status labels only. We decoded each of the cue condition combinations (e.g., non-target vs. target) separately to avoid that a single condition would entirely drive decoding performance for all four status conditions (i.e., one vs. all other three conditions).

For all decoding schemes, supervised learning models were used by training ridge regression classifiers to distinguish the output values (i.e., model targets) using scikit-learn's ridge algorithm (Pedregosa et al., 2011). Ridge regression is a regression model that uses linear least squares with L2 regularization. In essence, it computes a weighted combination of voxel activity values and by adjusting the per-voxel regression weights it minimizes the discrepancy between the predicted output value and the correct output value. By imposing a penalty on the size of the coefficients, this method is particularly robust against multicollinearity which is a phenomenon that is commonly observed in multi-voxel patterns and would otherwise lead to high susceptibility of random error in classic OLS methods.

For within-condition category decoding as well as status decoding, we evaluated classification performance with a standard Leave-One-Out cross-validation strategy. In other words, the classifier was always trained to learn the mapping between the neural patterns and the corresponding category (within-condition decoding) or status labels (status decoding) for all but one run. Then, the classifier was used to predict the category or status labels in the remaining, for this iteration independent, run. Since this procedure was repeated for each combination of runs, we ended up with eight iterations across which we calculated the average to obtain the overall classification performance. For cross-condition category decoding, the data of the tested condition was always independent and we could therefore train and test on all runs.

We confined our statistical analyses on the average classification performance across nine TRs following the analysis approach of van Loon and colleagues (Olmos-Solis et al., 2021; van Loon et al., 2018 and similar to Lee et al., 2013). Note that since our TR differed from these previous studies, we selected the TRs that most closely approximated their time interval which corresponded to time intervals of 6.3 secs in the present study. More specifically, we averaged the classification performance across each of two separate time intervals which corresponded to the fourth to the twelfth TR after cue (cue period) and search display onset (search period). Classification performance was compared against chance level (within- and cross-condition decoding: one over three or 33.33% due to the three object categories; status decoding: one over two or 50% due to the comparison between two status conditions) using one-sample *t*-tests (two-tailed). For each condition or combination comparison paired samples *t*-tests were used (two-tailed). If the normality assumption was violated (Shapiro–Wilks test), we also report Wilcoxon signed-rank tests. Following a reviewer's suggestion we also re-ran our analyses using randomizations tests instead of *t*-tests. This generated a very similar pattern of results and no different conclusions. We therefore stuck to our original analysis plan.

Representational dissimilarity analysis

To explore the representational geometry of the brain activity patterns in the different conditions, we created a representation dissimilarity matrix (RDM) for each TR and each participant (Kriegeskorte et al., 2008; Kriegeskorte & Kievit, 2013). Each value in such an RDM represents the dissimilarity between the activity patterns that are associated with two different exemplars (except along the matrix diagonal where it is associated with the same exemplars and dissimilarity is per definition zero). As the distance measure, we employed 1-rho (i.e., Spearman's rank correlation coefficient) across all voxels of the given ROI. For visualization purposes, we further replaced each element of the RDM by its rank in the distribution of all of the RDM's elements (scaled from 0 to 1). Thus, the RDM is a 48×48 matrix which represents the ranked pattern dissimilarity among all combinations of the four exemplars of each of the three categories (cows, dressers, skate) and the four different status levels (drop, non-target, neutral, target). Individual RDMs were averaged across all runs and TRs from the specific interval of interest (i.e., cue or search period).

Results

Behaviour

To assess whether the status cues and their associated templates were used to aid search performance, we submitted RT and response accuracy data to separate ANOVAs with cue status as factor (drop, neutral, non-target, target) for data from the behavioural and fMRI session separately. For the RT analyses all incorrect responses were excluded. The ANOVA for the RT in the behavioural session showed a significant main effect ($F(3, 63) = 37.584$, $p < .001$, $\eta p^2 = .642$). Figure 1(B) shows that participants searched faster with a target or a non-target cue relative to any of the two baseline conditions (target vs. drop: $t(21) = 6.59$, $p < .001$, Cohen's $d = 1.41$, neutral vs. target: $t(21) = 10.14$, $p < .001$, $d = 2.16$, non-target vs. drop: $t(21) = 3.57$, $p = .002$, $d = .76$, neutral vs. non-target: $t(21) = 6.34$, $p < .001$, $d = 1.352$). They were also faster in the drop condition than in the neutral cue condition ($t(21) = 3.22$, $p = .004$, $d = .69$). The relative slow responses in the latter condition were likely caused by the participants processing another cue at search onset (i.e., indicating what to do with the up to then neutral cue). Importantly, participants searched significantly faster with a target than a non-target cue (Wilcoxon $Z = 233$, $p < .001$, $r_b = .84$; $t(21) = 3.80$, $p = 0.001$, $d = .81$). The ANOVA on accuracy data was significant ($F(3, 63) = 2.854$, $p = .044$, $\eta p^2 = .120$). While accuracy was high (> 93.6%) in all conditions, paired samples *t*-tests revealed that participants were somewhat more accurate on trials with a drop cue compared to trials with a non-target cue (96.4%

vs. 93.7%, Wilcoxon $Z = 96$, $p = .043$, $r_b = .60$; $t(21) = 2.497$, $p = 0.021$, $d = .532$), with the other conditions falling in between.

Similar behavioural results were observed for the fMRI session (see Figure 1(C)). The ANOVA on RT showed a significant main effect ($F(3, 63) = 30.52$, $p < .001$, $\eta p^2 = .592$). Like in the behavioural session, participants searched faster with a target or non-target cue relative to any of the two baseline conditions (target vs. drop: $t(21) = 6.866$, $p < .001$, $d = 1.09$, neutral vs. target: $t(21) = 8.809$, $p < .001$, $d = 1.88$, non-target vs. drop: $t(21) = 5.132$, $p < .001$, $d = 1.09$, neutral vs. non-target: $t(21) = 5.497$, $p < .001$, $d = 1.17$). This time there was no significant difference in search speed between both baseline conditions, drop and neutral, during the fMRI session ($t(21) = 0.162$, $p = .873$, $d = .03$). Most importantly though, we also replicated the finding that search times were faster when participants searched with a target relative to a non-target cue ($t(21) = 3.220$, $p = .004$, $d = .687$). The ANOVA on accuracy data was significant as well ($F(3, 63) = 3.000$, $p = .029$, $\eta p^2 = .133$). Similar to the behavioural session, accuracy was high in all conditions during the fMRI session (> 93.1%) but participants gave slightly more correct responses in the drop cue relative to the neutral cue condition (95.7% vs. 93.2%, $t(21) = 2.783$, $p = .011$, $d = .593$) as well as in the target cue relative to the neutral cue condition (95.4% vs. 93.2%, $t(21) = 2.282$, $p = .033$, $d = .487$).

The behavioural results from both sessions indicate that participants used the informative cues during the search, whether such cues indicated targets or nontargets. As a consequence, search performance benefited relative to baseline, with target templates being more effective (> 200 ms faster on average) than non-target templates (> 131 ms faster on average), consistent with earlier findings. At the same time, target cues appeared more effective than non-target cues, which is also consistent with earlier reports.

fMRI

Overall, the fMRI analyses were guided by two main questions. First, which brain areas carry information about the *content* of the template (i.e., its category membership), and/or the *status* of the template (i.e., whether it is a target, non-target, neutral or dropped template). Second, if a brain area represents the content of the template, does it represent it in the *same* or a *different* way for the different types of templates (i.e., target, non-target, neutral or drop). For this reason, we focused on decoding analyses from the delay period after cue onset. Decoding analyses of the search period can be found in the Supplementary Materials (Supplementary Text) together with the univariate BOLD responses. Figures 2 and 5 show the decoding accuracies during the delay period for each of the three classification analyses across the different ROIs for each cue condition separately (see Supplementary Materials Figures 5–7 for the entire time courses for these and all following analyses).

Content decoding

To assess which brain areas represent category information of the memorized item (i.e., cow, dresser, or skate), classification accuracies from the *within-condition* category decoding scheme were submitted to an ANOVA with ROI (VC, LOC, pF, IPS, FEF and lPFC) and cue condition (target, non-target, neutral and drop) as factors. This analysis only showed a significant main effect of ROI ($F(5, 105) = 10.679$, $p < .001$, $\eta p^2 = .337$). Both the effect of cue condition ($F(3, 63) = .616$, $p = .607$, $\eta p^2 = .028$) and the interaction between ROI and cue condition failed to reach significance ($F(15, 315) = .507$, $p = .824$, $\eta p^2 = .024$). As is evident from Figure 2, within-condition category decoding was reliably above chance for areas VC, LOC and pF but not for IPS, FEF and lPFC. Thus, the template's content was represented predominantly in occipito-temporal areas. Importantly, in none of the ROIs was the decoding of non-target templates weaker than of target templates, or than of neutral templates (all $p > .1$).

To further investigate whether the neural representations of the templates in the different cue conditions were similar or different, we inspected the RDMs for each of the ROIs during the delay period. Figure 3 strongly suggests that the representational geometry of the activity patterns in the drop condition differed significantly from the other three conditions. In the context of the task instructions, the drop condition deviated from the other conditions in that it was the only condition where participants could altogether drop the object from memory from the start of the delay period. Given that there was also no indication that non-target templates were more weakly represented than dropped templates (Figure 2), all further analyses were done without

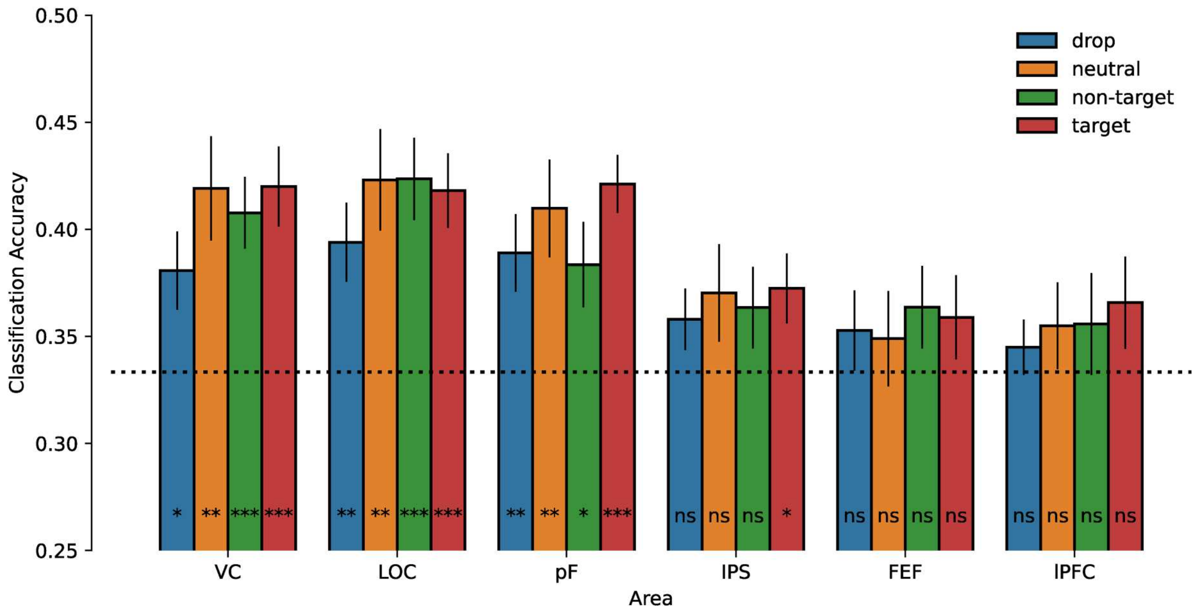

Figure 2. Object category decoding during the delay period (cross-condition) for areas VC, LOC, pF, IPS, FEF and lPFC separately for each cue condition. Significance levels against chance here and in all other figures are indicated by asterisks whereby † < .1, * $p < .05$, ** $p < .01$, *** $p < .001$, *ns*: not significant.

the drop conditions. Figure 3 suggests, at first glance, that the remaining three conditions are very similar, regardless of whether a target, non-target or neutral cue was shown. In general, then, participants used the cues to store objects in memory when useful for the search task (see Supplementary Materials Figure 2 for the RDM of the search period and Figures 3 and 4 for the RDMs for both periods but without the drop condition).

To quantify and statistically test these observations, we calculated the average similarity of the neural representations within and between the cue conditions. For the *within-condition* similarity condition, we computed the average correlation of each exemplar of each category with all the other exemplars of that category for each cue condition (e.g., the correlation between one exemplar of a cow and another exemplar of a cow in the target condition). For the *between-condition* similarity condition, we calculated the average correlation of each exemplar of each category in one condition with all the other exemplars of that same category but in another condition (e.g., the correlation between one exemplar of a cow in the target condition and another exemplar of a cow in the non-target condition). The resulting ANOVA with similarity type (within- vs. between-condition), ROI (VC, LOC, pF, IPS, FEF and lPFC) and cue condition (neutral vs. non-target, neutral vs. target and non-target vs. target), showed a significant main effect of similarity type ($F(1, 21) = 7.047$, $p = .015$, $\eta p^2 = .251$), ROI ($F(5, 105) = 13.269$, $p < .001$, $\eta p^2 = .387$), but not cue condition ($F(2, 42) = .054$, $p = .873$, $\eta p^2 = .003$) or any of the interactions (all $F < .499$, $p > .586$, $\eta p^2 < .023$). Figure 4 shows the pattern of correlations for the crucial non-target vs. target cue conditions. Importantly, correlations were of similar magnitude and followed the same pattern across brain areas for the between-condition and within-condition comparisons, suggesting a similar type of representation across target and non-target conditions.

To further address the question of whether the content of a template was represented in the same or a different way for the different types of templates, we analysed the data from the *cross-condition* category decoding scheme (see Figure 5(A)). An ANOVA with ROI (VC, LOC, pF, IPS, FEF and lPFC) and cue condition (neutral vs. non-target, non-target vs. target, neutral vs. target) on classification accuracies showed a main effect of ROI ($F(5, 105) = 13.008$, $p < .001$, $\eta p^2 = .382$), no main effect for condition cue ($F(2, 42) = .707$, $p = .499$, $\eta p^2 = .033$) and no interaction ($F(10, 210) = 1.515$, $p = .109$, $\eta p^2 = .080$). Classification accuracy for the non-target vs. target pair deviated significantly from chance in all ROIs (VC: $t(21) = 4.224$, $p < .001$, $d = .900$, LOC: $t(21) =$

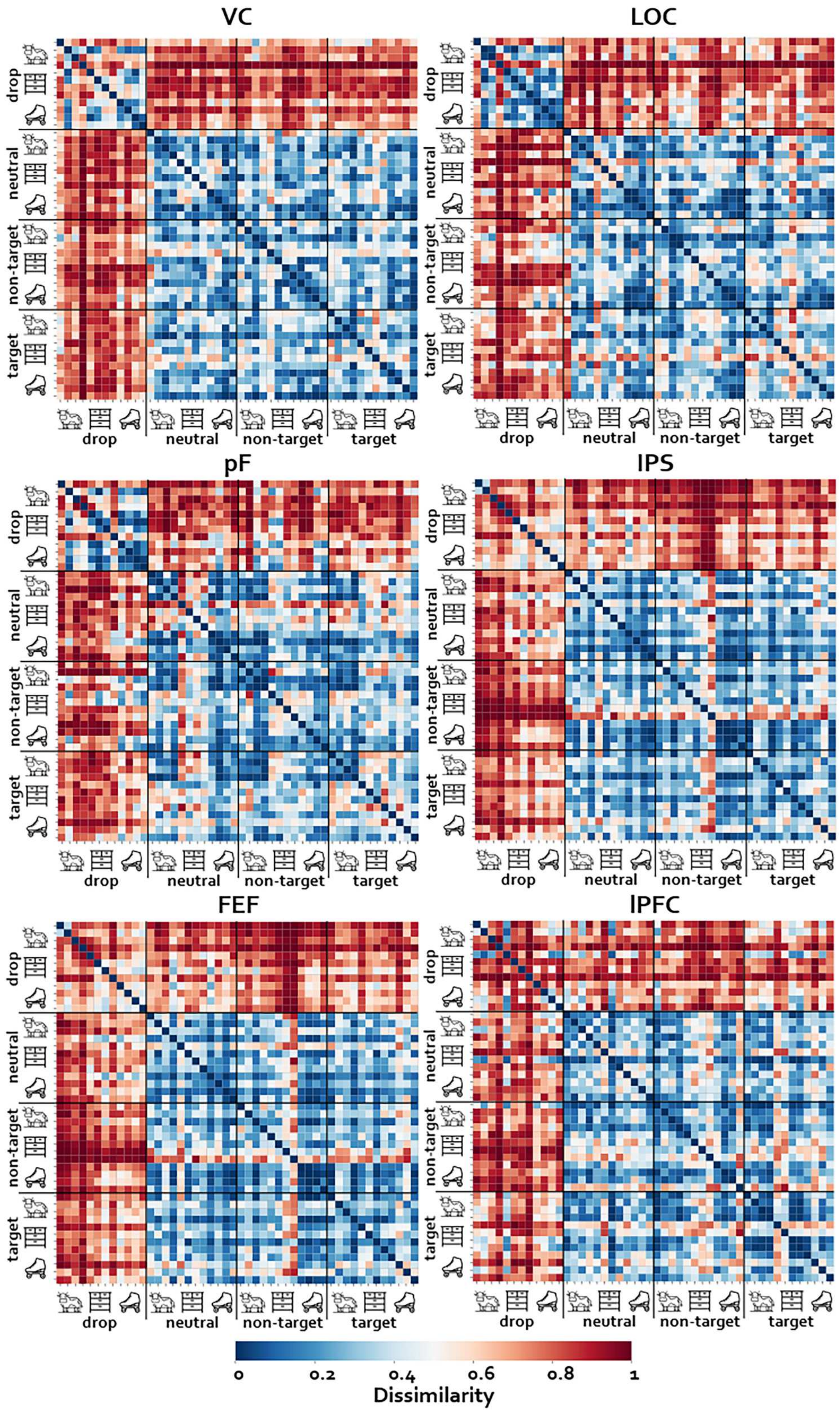

Figure 3. Representational dissimilarity analysis of object representations in areas VC, LOC, pF, IPS, FEF and lPFC during the delay period. The bluer the more similar and the redder the more dissimilar a representation. *Note:* Here and in all other representational dissimilarity graphs, dissimilarity is represented by scaled-ranked correlations.

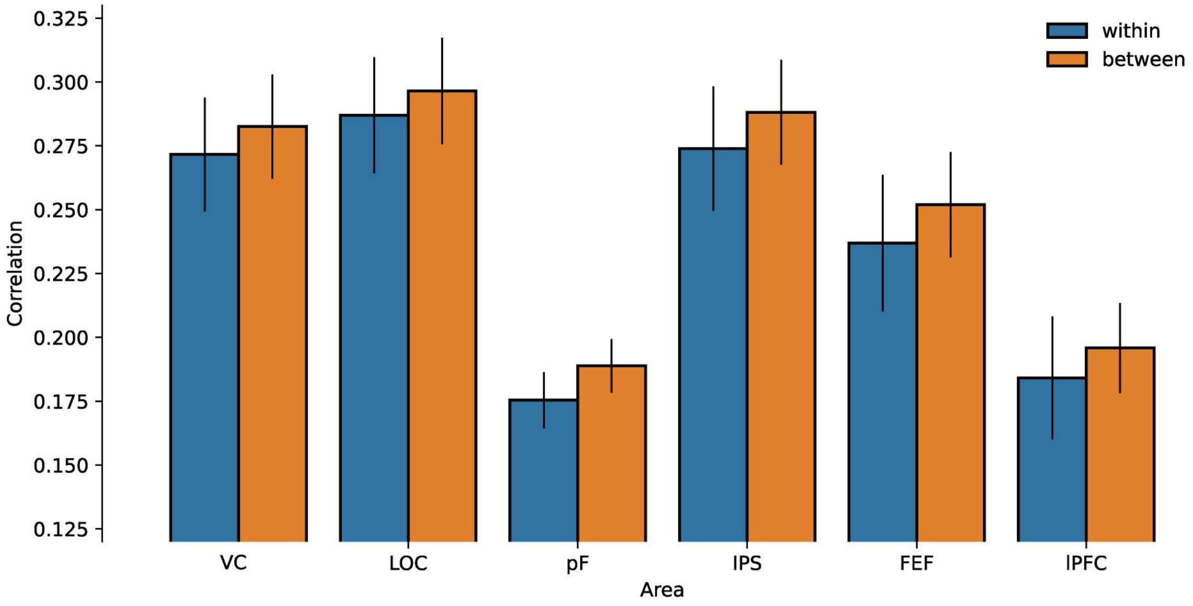

Figure 4. Correlations between representations within and between target and nontarget template conditions, for areas VC, LOC, pF, IPS, FEF and lPFC in the non-target vs. target cue condition.

4.241, $p < .001$, $d = .904$, pF: $t(21) = 3.988$, $p < .001$, $d = .850$, IPS: $t(21) = 4.024$, $p < .001$, $d = .858$, FEF: $t(21) = 2.846$, $p = .010$, $d = .607$) except in lPFC (lPFC: $t(21) = 1.711$, $p = .102$, $d = .365$). These results further confirm little difference in neural representations of the same object category when used as a target versus non-target cue.

Status decoding

The previous analyses were concerned with the representation of the *content* of the template. Here, we used the *status* decoding scheme to investigate whether different brain regions differentiate between the relevance status of a template, *regardless* of its content (see Figure 5(B)). An ANOVA with ROI (VC, LOC, pF, IPS, FEF and lPFC) and cue condition (neutral vs. non-target, non-target vs. target, neutral vs. target) showed a main effect of ROI ($F(5, 105) = 3.187$, $p = .010$, $\eta p^2 = .132$), no main effect of cue condition ($F(2, 42) = 2.103$, $p = .135$, $\eta p^2 = .091$) and no interaction ($F(10, 210) = .876$, $p = .507$, $\eta p^2 = .040$). Reliable above-chance decoding for all condition pairs within a particular ROI was only observed in lPFC. Critical for our main hypothesis was that the same pattern was even more evident in the data of the non-target vs. target pair ($t(105) = 2.366$, $p = .020$). Here, significant above-chance decoding of status information was only found in more frontal areas FEF ($t(21) = 2.091$, $p = .049$, $d = .446$) and lPFC ($t(21) = 3.208$, $p = .004$, $d = .684$). The comparisons in the other ROIs did not reach significance (VC: $t(21) = 2.014$, $p = .057$, $d = .429$, LOC: $t(21) = .798$, $p = .434$, $d = .170$, pF: $t(21) = .201$, $p = .843$, $d = .043$, IPS: $t(21) = 2.054$, $p = .053$, $d = .438$). Notably, there was a significantly larger univariate BOLD response for the non-target relative to the target cue condition in area lPFC ($t(21) = 3.032$, $p = .006$, $d = .646$; Supplementary Figure 1A). None of the other comparisons showed a significant difference in the BOLD response. It is likely that the difference in overall BOLD response contributed to above-chance status decoding for the non-target vs. target pair. However, that does not necessarily mean it was the only contribution, particularly given the fact that the other comparisons also resulted in above-chance multivariate decoding while there was no univariate difference.

Conclusion

Overall, these results suggest that while posterior regions (i.e., VC, LOC and pF) predominantly represent object category-specific *content* information of the template, frontal areas (i.e., FEF and lPFC) predominantly carry information about the *status* of the template, with IPS exhibiting a tendency to represent both to some extent. Figure 6 illustrates this pattern more clearly, by combining content and status decoding for the target and non-target conditions in one

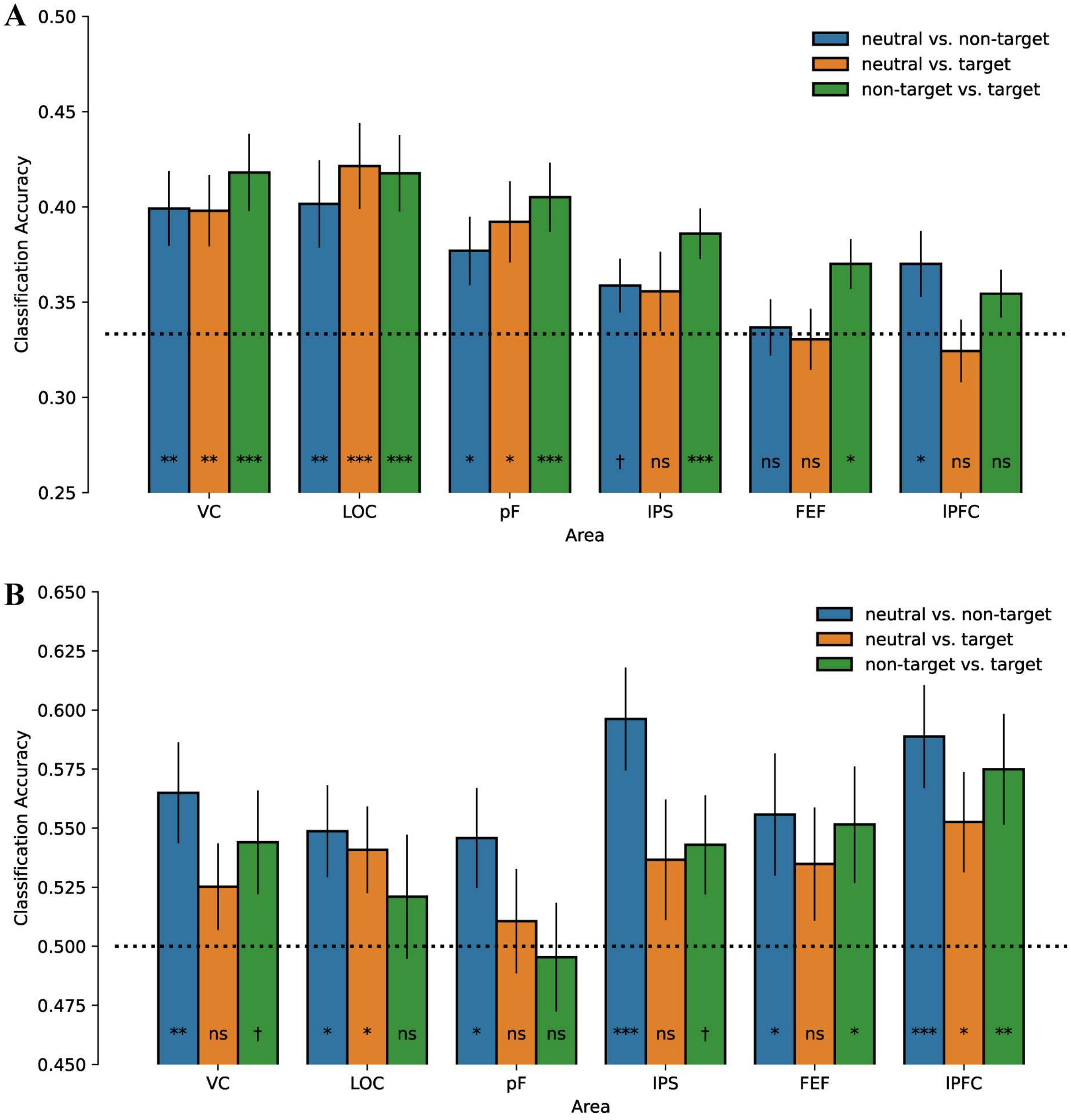

Figure 5. Object category and status decoding during the delay period for areas VC, LOC, pF, IPS, FEF and lPFC. (A) Classification accuracy for cross-condition category decoding separately for each cue condition. (B) Classification accuracy for status decoding separately for each cue condition. Significance levels against chance here and in all other figures are indicated by asterisks whereby † < .1, * $p < .05$, ** $p < .01$, *** $p < .001$, *ns*: not significant.

graph. A two-way ANOVA on the decoding accuracies from target and non-target trials with ROI (VC, LOC, pF, IPS, FEF and lPFC) and decoding scheme (category content and status) as factors indeed showed a significant interaction ($F(5, 105) = 8.036$, $p < .001$, $\eta p^2 = .277$) providing statistical support for the pattern observed. We point out though that somewhat unexpectedly, early visual cortex (VC) also showed substantial above-chance status decoding – something we will return to in the Discussion.

Discussion

Consistent with previous work (e.g., Arita et al., 2012; Carlisle & Nitka, 2019; Woodman & Luck, 2007; Zhang et al., 2020), we found that providing participants with information on which objects can be safely ignored improves visual search performance, as RTs were faster when participants were given a non-target cue (as to what to ignore) compared to when no cue or a less informative neutral cue was provided.

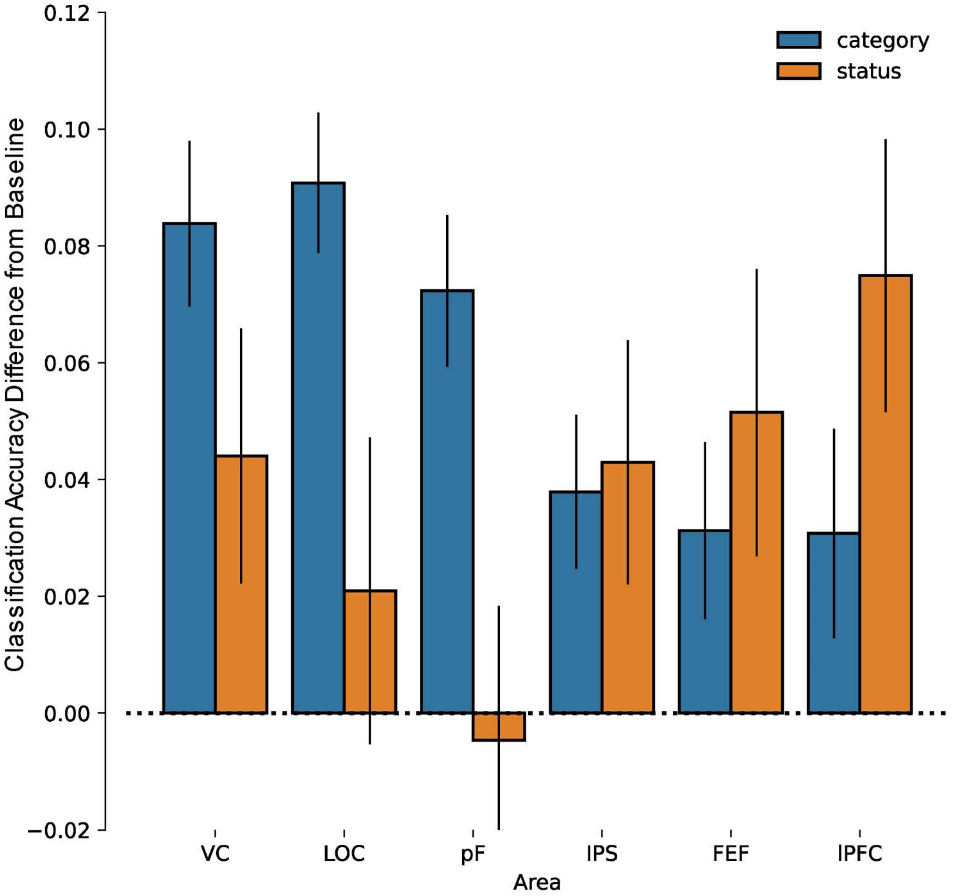

Figure 6. Direct comparison of cross-condition category versus status decoding for target and non-target cue conditions during the delay period. Classification accuracy is shown collapsed across the different cue conditions and shown relative to the relative chance level.

At the same time, as in these previous studies, non-target cues were not as effective as target cues. The reduced efficacy of non-target cues compared to target cues could reflect several things. For one, inhibition of distractors, whether implemented prior to display onset or not, might be less effective than enhancement of potential targets. Alternatively, non-target cues may be used to derive the target identity at the moment the display appears – a process that takes additional time (Beck & Hollingworth, 2015; Becker et al., 2015). Finally, participants may try to prevent inadvertent orienting to cued distractors by slowing down ("gating") the response to the overall visual input (Noonan et al., 2018; de Vries et al., 2019). However, our main question was not what makes non-target cues less effective than target cues, but what makes them effective in the first place.

We distinguished two main hypotheses. Under the first hypothesis, the representation of non-target information is either suppressed or transformed already prior to search so that it is effectively deprioritized by the time the non-target information actually appears. We found no evidence supporting this hypothesis. Specifically, using MVPA, we found no evidence that the representational pattern of activity in object-selective cortex was weaker for non-target templates than for target templates or for memories in the neutral baseline condition – let alone that activity would be weaker than in the drop condition, where the memory was no longer relevant. An analysis of representational similarity also did not reveal any differences between target, non-target or neutral templates, and there was no evidence that non-target representations were transformed in ways that have been found in previous studies (van Loon et al., 2018; Wan et al., 2020; Yu et al., 2020). Taken together, this provides little evidence in support of the inhibition hypothesis.

Instead, the results of our study are consistent with the alternative hypothesis, namely that target and non-target information is represented similarly prior to search, but accompanied by the implementation of different *control signals* that prepare for the right course of action when the search display, and with it the target or non-target information, is encountered. While content representations in the posterior cortex remained the same across conditions, frontal cortical areas were clearly sensitive to the status of the template per se, regardless of content. This was the case for FEF, but especially for lPFC, which could clearly distinguish status, but not category. Contrast this with pF, which could clearly distinguish category, but was not sensitive to the difference between target and non-target status. Currently, we have no way of determining the exact nature of the frontal operations that are induced by the different cues. We show that these areas are sensitive to what we have referred to as the status of the cued item, where status is not intended as a neural property in and of itself but as a neutral operational term meant to convey the prospective functional use of the stored representation for the subsequent search (cf. van Loon et al., 2017, Olmos-Solis et al., 2021). Many potential processes are likely to be involved in realizing that use. As laid out in the introduction, this could involve a plan to reactively suppress the distractor once it is actually encountered (cf. Moher and Egeth, 2012; Gaspelin & Luck, 2018), an overall suppressive signal gating the perceptual input at search onset (cf. Noonan et al., 2018; de Vries et al., 2019), or a plan to recode the non-target information into target information once the display has appeared (cf. Beck & Hollingworth, 2015; Becker et al., 2015). The successful status decoding may also reflect differences in the overall effort or task difficulty that comes with any of these operations, and that would then presumably result in different energy consumption and thus blood oxygenation patterns. If so, this would still need an explanation for why one condition is more effortful than the other, and we believe here the potential additional cognitive control operations that may come with non-target templates would be a good candidate. In short, any or all of these aspects might have contributed, as may be effects as yet unthought of, and future research will need to disentangle these contributions. Regardless, we believe the concept of a "negative" or "rejection" template (cf. Arita et al., 2012; Woodman & Luck, 2007) still has value, but then reflect how search will be controlled, rather than how the content representation of the template itself changes.

Here we focused on the flexible prior inhibition of distractor object features or identities. Distractor inhibition may also build up over many trials in which the same information is repeated over and over again. In addition, inhibition may operate spatially, by suppressing a known distractor location. These potential mechanisms are treated elsewhere (e.g., Chelazzi et al., 2019; Cunningham & Egeth, 2016; Failing et al., 2019; van Moorselaar & Slagter, 2019; van Moorselaar et al., 2020; 2021; van Zoest et al., 2021; Wang et al., 2019). In general, observers may effectively learn to avoid a consistent distractor (whether in terms of location or feature or both). Given the component of learning, such avoidance mechanisms must reflect traces or mechanisms that survive beyond the current trial and are thus by definition present prior to the next search display onset. However, while that may count as prior inhibition (although such avoidance mechanisms are not necessarily inhibitory in nature), they certainly would not be flexible in the sense that observers can instantiate the suppression on a trial-by-trial basis.

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, consistent with lPFC and FEF representing the goal- or task context-related signals that prioritize relevant over irrelevant information (e.g., Christophel et al., 2018; Fuster et al., 1985; Lee & D'Esposito, 2012; Lee et al., 2013; Miller et al., 2011; Serences, 2016; Zanto et al., 2011). Note that the opposite trends for category and status decoding across the cortical hierarchy preclude an explanation in terms of mere differences in signal to noise levels across regions, and must thus represent different functionalities. The results are also consistent with a recent study by Olmos-Solis et al. (2021), who showed a similar division of labor in a study in which the relevance of the working memory representation was not determined by whether it would be a target or non-target, but by the specific order of tasks. That is, observers were asked to remember a target cue, which would then be relevant for either the first or the second of two sequential visual search tasks. While posterior areas (early visual cortex and pF) were more sensitive to the target category than to its status, frontal areas (including FEF and lPFC) were sensitive to the order of relevance for the upcoming tasks, irrespective of category. In addition, Olmos-Solis et al. found the IPS to be sensitive to both target category, and target status, suggesting that IPS served as a hub multiplexing target information with task rules such as those setting target priority (cf. Cole et al., 2013; Majerus et al., 2007; Mitchell & Cusack, 2008; Sprague et al., 2018). However, in the current study, we found IPS to not be particularly sensitive to the status of the cue (i.e., whether target or non-target), nor the cue category. This suggests that IPS was not particularly active or even involved in the a priori (de-)prioritization of information in this task.

A number of previous fMRI studies have attempted to find evidence for distractor suppression. Seidl et al. (2012) asked participants to search pictures of real-world scenes for a specific target category, such as a car, a human, or a tree. The target category was consistent per run (block). Crucially, the target of a previous run could return as a distractor in a current run. If so, this led to suppressed BOLD response in distractor-related areas in object-selective cortex (see also Peters et al., 2012). However, activity was measured after scene onset and could thus represent pro-active or reactive mechanisms. Moreover, given the blocked nature of the task, the pattern might represent a gradual build-up of suppression, rather than in a flexible top-down fashion. The same was true for a study by Marini et al. (2016), who used a paradigm in which participants were instructed to respond to a target that could be surrounded by distractors, which in turn could be congruent or incongruent. The probability of distractors being present as well as of their congruency was varied between blocks. Marini et al found reduced activity in the occipital cortex in response to display onsets when incongruent distractors were expected, suggestive of suppression. They also found lPFC to be involved in anticipating distraction. However, here too the design did not allow for a test of flexible, goal-driven suppression prior to search. Finally, using a similar design as ours but with colours rather than object categories, Reeder et al. (2017, 2018) found an overall reduced BOLD response, as well as fuzzier (i.e., less consistent) representational patterns in the occipital cortex for non-target cues than for target cues. Here we did not replicate this finding. It is possible that participants indeed partially suppressed the non-target representation in the case of colour, and for an unknown reason did not do so in our experiment. Conversely, it is also possible that participants found the non-target cue not very useful, and simply devoted less memory to it than to target cues. Alternatively, participants may have verbally labelled the non-target colour, thus reducing the necessity for a visual memory. In the current study, verbal re-labelling was discouraged by using different exemplars from the same category. In any case, it deserves emphasizing that although here we found no evidence for active prior suppression, we of course cannot exclude the possibility that such suppression does occur under different circumstances.

Limitations

A number of limitations to our study can be pointed out. For one, the search displays consisted of exemplars within the same category (e.g., two types of cow, one set containing the target, the other set only nontargets). This may mean that possibility for attentional guidance was quite limited, and hence any of the behavioural advantages due to the informative cues (whether target or non-target cues) was not due to more efficient attentional guidance but due to post-selection decision processes. We cannot exclude the possibility of at least a partial contribution of such decision processes. This would mean that any templates tested were not necessarily *attentional* templates, used as a pro-active attentional set (cf. Olivers et al., 2011). In our view, they would still be *target* templates though, as also decision processes will need some form of target or non-target representation. We show that such representations are no different after target or non-target cues.

Note further that we chose to employ multiple similar exemplars for each of the categories as a basis for the search display in order to prevent straightforward recoding of non-target information into target information (cf. Beck & Hollingworth, 2015; Becker et al., 2015). If the target was not presented on one cow, it would be presented on any of three other cows. While we consider such conversion unlikely due to the additional memory load and therefore effort that it would entail, neither can we fully exclude the possibility. Such a conversion would explain the similarity between the target and non-target template conditions. In any case, it would imply that the observer prefers such additional load over some alternative, such as creating an inhibitory template. We point out that the same logic in principle also applies to target cues: Observers could convert that information into distractor information instead. While even less likely, we cannot logically exclude this possibility.

Finally, there is the possibility that observers treated *all* pictures of objects as distractors. Note that the actual search target (the letter T among Ls) was printed on top of the objects, and so the objects may actually compete with selection through suppressive interactions (e.g., Kastner et al., 2001). It would then be wise to inhibit all pictures regardless of being indicative of targets or nontargets – which would then lead to highly similar representations. While we cannot exclude this possibility, the behavioural data makes it an unlikely contender: The cued information was useful for search, and this was more so the case for target than for non-target cues.

Another potential limitation is the consistent mapping of cue appearance (+, –, O, X) to meaning (positive or target template, negative or non-target template, neutral, drop). We opted for this in order to maximize understanding and perceived usefulness of the cue (compared to more arbitrary and abstract cue-meaning mappings), but it runs the risk that we have been decoding the sensory appearance of the cue rather than the cognitive state it was meant to evoke. We regard this risk as low for the higher-order frontal and parietal areas which are traditionally more sensitive to task demands than to sensory input. Moreover, the areas maximally sensitive to perceptual category (LOC, pF) were actually least sensitive to status. However, the visual differences between cues may well explain the rather prominent status decoding in early visual cortex that we observed.

Conclusion

To conclude, so far neurophysiological measures have found little direct support for flexible, pro-active, a priori suppression of distractor representations (see also de Vries et al., 2019; Noonan et al., 2018; Rajsic et al., 2020; van Zoest et al., 2021 for findings from the EEG domain). Instead, the system appears to plan how to react when the stimulus appears.

Note

1. Here, we focus on the *flexible* prior suppression of distractor features or identities. Distractor suppression may also build up over many trials in which the same information is repeated from trial to trial. In addition, suppression may operate spatially, by inhibiting a known distractor location. We will return to this in the Discussion.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the H2020 European Research Council (grant number ERC-2013-CoG-615423) and Nederlandse Organisatie voor Wetenschappelijk Onderzoek (grant number 453-16-002).

ORCID

Stefan Pollmann http://orcid.org/0000-0001-5840-5658
Christian N. L. Olivers http://orcid.org/0000-0001-7470-5378

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*, 14. https://doi.org/10.3389/fninf.2014.00014

Arita, J. T., Carlisle, N. B., & Woodman, G. F. (2012). Templates for rejection: Configuring attention to ignore task-irrelevant features. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 580. https://doi.org/10.1037/a0027885

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Baeck, A., Wagemans, J., & Beeck, H. P. O. (2013). The distributed representation of random and meaningful object pairs in human occipitotemporal cortex: The weighted average as a general rule. *Neuroimage*, *70*, 37–47.

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*(5), 1195–1205.

Battistoni, E., Stein, T., & Peelen, M. V. (2017). Preparatory attention in visual cortex. *Annals of the New York Academy of Sciences*, *1396*(1), 92–107. https://doi.org/10.1111/nyas.13320

Beck, V. M., & Hollingworth, A. (2015). Evidence for non-target feature guidance in visual search is explained by spatial recoding. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1190. https://doi.org/10.1037/xhp0000109

Becker, M. W., Hemsteger, S., & Peltier, C. (2015). No templates for rejection: A failure to configure attention to ignore task-irrelevant features. *Visual Cognition*, *23*(9-10), 1150–1167. https://doi.org/10.1080/13506285.2016.1149532

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Bettencourt, K. C., & Xu, Y. (2016). Decoding the content of visual short-term memory under distraction in occipital and parietal areas. *Nature Neuroscience*, *19*(1), 150–157.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523.

Carlisle, N. B., & Nitka, A. W. (2019). Location-based explanations do not account for active attentional suppression. *Visual Cognition*, *27*(3–4), 305–316. https://doi.org/10.1080/13506285.2018.1553222

Chawla, D., Rees, G., & Friston, K. J. (1999). The physiological basis of attentional modulation in extrastriate visual areas. *Nature Neuroscience*, *2*(7), 671–676. https://doi.org/10.1038/10230

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology*, *80*(6), 2918–2940. https://doi.org/10.1152/jn.1998.80.6.2918

Chelazzi, L., Marini, F., Pascucci, D., & Turatto, M. (2019). Getting rid of visual distractors: The why, when, how, and where. *Current Opinion in Psychology*, *29*, 135–147. https://doi.org/10.1016/j.copsyc.2019.02.004

Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, *363*(6427), 345–347. https://doi.org/10.1038/363345a0

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. https://doi.org/10.1038/s41593-018-0094-4

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*(9), 1348–1355.

Conci, M., Deichsel, C., Müller, H. J., & Töllner, T. (2019). Feature guidance by negative attentional templates depends on search difficulty. *Visual Cognition*, *27*(3–4), 317–326. https://doi.org/10.1080/13506285.2019.1581316

Cosman, J. D., Lowe, K. A., Zinke, W., Woodman, G. F., & Schall, J. D. (2018). Prefrontal control of visual distraction. *Current Biology*, *28*(3), 414–420. https://doi.org/10.1016/j.cub.2017.12.023

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014

Cunningham, C. A., & Egeth, H. E. (2016). Taming the white bear: Initial costs and eventual benefits of distractor inhibition. *Psychological Science*, *27*(4), 476–485. https://doi.org/10.1177/0956797615626564

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

de Hollander, G., & Knapen, T. (2018, October 16). VU-Cog-Sci/nideconv:First alpha version nideconv (Version v0.1.0) Zenodo. https://doi.org/10.5281/zenodo.1463839

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

de Vries, I. E., Savran, E., van Driel, J., & Olivers, C. N. (2019). Oscillatory mechanisms of preparing for visual distraction. *Journal of Cognitive Neuroscience*, *31*(12), 1873–1894. https://doi.org/10.1162/jocn_a_01460

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Ilkay Isik, A., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Failing, M., Feldmann-Wüstefeld, T., Wang, B., Olivers, C., & Theeuwes, J. (2019). Statistical regularities induce spatial as well as feature-specific suppression. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(10), 1291. https://doi.org/10.1037/xhp0000660

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030. https://doi.org/10.1037/0096-1523.18.4.1030

Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., & Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

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.

Gaspelin, N., & Luck, S. J. (2018). The role of inhibition in avoiding distraction by salient stimuli. *Trends in Cognitive Sciences*, *22*(1), 79–92. https://doi.org/10.1016/j.tics.2017.11.001

Gayet, S., Guggenmos, M., Christophel, T. B., Haynes, J.-D., Paffen, C. L. E., Van der Stigchel, S., & Sterzer, P. (2017). Visual working memory enhances the neural response to matching visual input. *Journal of Neuroscience*, *37*(28), 6638–6647. https://doi.org/10.1523/JNEUROSCI.3418-16.2017

Giesbrecht, B., Weissman, D. H., Woldorff, M. G., & Mangun, G. R. (2006). Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks. *Brain Research*, *1080*(1), 63–72. https://doi.org/10.1016/j.brainres.2005.09.068

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, *5*, 13. https://doi.org/10.3389/fninf.2011.00013

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10-11), 1409–1422.

Grubert, A., Fahrenfort, J., Olivers, C. N., & Eimer, M. (2017). Rapid top-down control over template-guided attention shifts to multiple objects. *Neuroimage*, *146*, 843–858. https://doi.org/10.1016/j.neuroimage.2016.08.039

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*, *111*(10), E962–E971. https://doi.org/10.1073/pnas.1312567111

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology*, *86*(3), 1398–1411. https://doi.org/10.1152/jn.2001.86.3.1398

Klein, A., Ghosh, S.S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E, & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLoS Computational Biology*, *13*(2), e1005350. https://doi.org/10.1371/journal.pcbi.1005350

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412. https://doi.org/10.1016/j.tics.2013.06.007

Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(1), 4. https://doi.org/10.3389/neuro.01.016.2008

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., & Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, *10*(3), 120–131.

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

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. https://doi.org/10.1523/JNEUROSCI.0627-12.2012

Lorenc, E. S., Sreenivasan, K. K., Nee, D. E., Vandenbroucke, A. R., & Esposito, M. (2018). Flexible coding of visual working memory representations during distraction. *Journal of Neuroscience*, *38*(23), 5267–5276.

MacLeod, C. M. (2007). The concept of inhibition in cognition. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 3–23). American Psychological Association.

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.

Mansouri, F. A., Koechlin, E., Rosa, M. G., & Buckley, M. J. (2017). Managing competing goals-a key role for the frontopolar cortex. *Nature Reviews Neuroscience*, *18*(11), 645–657.

Marini, F., Demeter, E., Roberts, K. C., Chelazzi, L., & Woldorff, M. G. (2016). Orchestrating proactive and reactive mechanisms for filtering distracting information: Brain-behavior relationships revealed by a mixed-design fMRI study. *Journal of Neuroscience*, *36*(3), 988–1000. https://doi.org/10.1523/JNEUROSCI.2966-15.2016

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324.

Miller, B. T., Vytlacil, J., Fegen, D., Pradhan, S., & Esposito, M. (2011). The prefrontal cortex modulates category selectivity in human extrastriate cortex. *Journal of Cognitive Neuroscience*, *23*(1), 1–10.

Mitchell, D. J., & Cusack, R. (2008). Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks. *Cerebral Cortex*, *18*(8), 1788–1798.

Moher, J., & Egeth, H. E. (2012). The ignoring paradox: Cueing distractor features leads first to selection, then to inhibition of to-be-ignored items. *Attention, Perception, & Psychophysics*, *74*, 1590–1605.

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage*, *59*(3), 2636–2643. https://doi.org/10.1016/j.neuroimage.2011.08.076

Noonan, M. P., Crittenden, B. M., Jensen, O., & Stokes, M. G. (2018). Selective inhibition of distracting input. *Behavioural Brain Research*, *355*, 36–47.

Olivers, C. N., & Eimer, M. (2011). On the difference between working memory and attentional set. *Neuropsychologia*, *49*(6), 1553–1558. https://doi.org/10.1016/j.neuropsychologia.2010.11.033

Olivers, C. N., 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), 327–334.

Olmos-Solis, K., van Loon, A., Los, S., & Olivers, C. (2018). Microsaccades reveal the temporal dynamics of template and response preparation during visual search. *Journal of Vision*, *18*(10), 1031–1031. https://doi.org/10.1167/18.10.1031

Olmos-Solis, K., van Loon, A. M., & Olivers, C. N. (2021). Content or status: Frontal and posterior cortical representations of object category and upcoming task goals in working memory. *Cortex*, *135*, 61–77. https://doi.org/10.1016/j.cortex.2020.11.011

Paus, T. (1996). Location and function of the human frontal eye-field: A selective review. *Neuropsychologia*, *34*(6), 475–483.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.

Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences*, *108*(29), 12125–12130. https://doi.org/10.1073/pnas.1101042108

Peters, J. C., Roelfsema, P. R., & Goebel, R. (2012). Task-relevant and accessory items in working memory have opposite effects on activity in extrastriate cortex. *Journal of Neuroscience*, *32*(47), 17003–17011. https://doi.org/10.1523/JNEUROSCI.0591-12.2012

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*, *84*, 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048

Prince, J. S., Charest, I., Kurzawski, J. W., Pyles, J. A., Tarr, M. J., & Kay, K. N. (2022). Improving the accuracy of single-trial fMRI response estimates using GLMsingle. *Elife*, *11*, e77599. https://doi.org/10.7554/eLife.77599

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. *Nature Neuroscience*, *22*(8), 1336–1344.

Rajsic, J., Carlisle, N. B., & Woodman, G. F. (2020). What not to look for: Electrophysiological evidence that searchers prefer positive templates. *Neuropsychologia*, *140*, 107376. https://doi.org/10.1016/j.neuropsychologia.2020.107376

Reeder, R. R., Olivers, C. N., Hanke, M., & Pollmann, S. (2018). No evidence for enhanced distractor template representation in early visual cortex. *Cortex*, *108*, 279–282.

Reeder, R. R., Olivers, C. N., & Pollmann, S. (2017). Cortical evidence for negative search templates. *Visual Cognition*, *25*(1–3), 278–290.

Sawaki, R., & Luck, S. J. (2013). Active suppression after involuntary capture of attention. *Psychonomic Bulletin & Review*, *20*(2), 296–301. https://doi.org/10.3758/s13423-012-0353-4

Seidl, K. N., Peelen, M. V., & Kastner, S. (2012). Neural evidence for distracter suppression during visual search in real-world scenes. *Journal of Neuroscience*, *32*(34), 11812–11819. https://doi.org/10.1523/JNEUROSCI.1693-12.2012

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research*, *128*, 53–67.

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.

Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences*, *106*(46), 19569–19574. https://doi.org/10.1073/pnas.0905306106

Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

van Loon, A. M., Olmos-Solis, K., Fahrenfort, J. J., & Olivers, C. N. (2018). Current and future goals are represented in opposite patterns in object-selective cortex. *ELife*, *7*, e38677. https://doi.org/10.7554/eLife.38677

van Loon, A. M., Olmos-Solis, K., & Olivers, C. N. (2017). Subtle eye movement metrics reveal task-relevant representations prior to visual search. *Journal of Vision*, *17*(6), 13–13. https://doi.org/10.1167/17.6.13

van Moorselaar, D., Daneshtalab, N., & Slagter, H. A. (2021). Neural mechanisms underlying distractor inhibition on the basis of feature and/or spatial expectations. *Cortex*, *137*, 232–250. https://doi.org/10.1016/j.cortex.2021.01.010

van Moorselaar, D., Lampers, E., Cordesius, E., & Slagter, H. A. (2020). Neural mechanisms underlying expectation-dependent inhibition of distracting information. *Elife*, *9*, e61048. https://doi.org/10.7554/eLife.61048

van Moorselaar, D., & Slagter, H. A. (2019). Learning what is irrelevant or relevant: Expectations facilitate distractor inhibition and target facilitation through distinct neural mechanisms. *Journal of Neuroscience*, *39*(35), 6953–6967. https://doi.org/10.1523/JNEUROSCI.0593-19.2019

van Zoest, W., Huber-Huber, C., Weaver, M. D., & Hickey, C. (2021). Strategic distractor suppression improves selective control in human vision. *Journal of Neuroscience*, *41*(33), 7120–7135. https://doi.org/10.1523/JNEUROSCI.0553-21.2021

Varoquaux, G., Gramfort, A., Pedregosa, F., Michel, V., & Thirion, B. (2011). Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. In . *In Information Processing in Medical Imaging: 22nd International Conference, IPMI 2011*, Kloster Irsee, Germany, July 3-8, 2011. Proceedings 22 (pp. 562–573). Springer Berlin Heidelberg.

Wan, Q., Cai, Y., Samaha, J., & Postle, B. R. (2020). Tracking stimulus representation across a 2-back visual working memory task. *Royal Society Open Science*, *7*(8), 190228. https://doi.org/10.1098/rsos.190228

Wang, B., van Driel, J., Ort, E., & Theeuwes, J. (2019). Anticipatory distractor suppression elicited by statistical regularities in visual search. *Journal of Cognitive Neuroscience*, *31*(10), 1535–1548. https://doi.org/10.1162/jocn_a_01433

Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search?. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(2), 363.

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. https://doi.org/10.1371/journal.pbio.3000769

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*(5), 656–661.

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57. https://doi.org/10.1109/42.906424

Zhang, Z., Gapelin, N., & Carlisle, N. B. (2020). Probing early attention following negative and positive templates. *Attention, Perception, & Psychophysics*, *82*(3), 1166–1175. https://doi.org/10.3758/s13414-019-01864-8