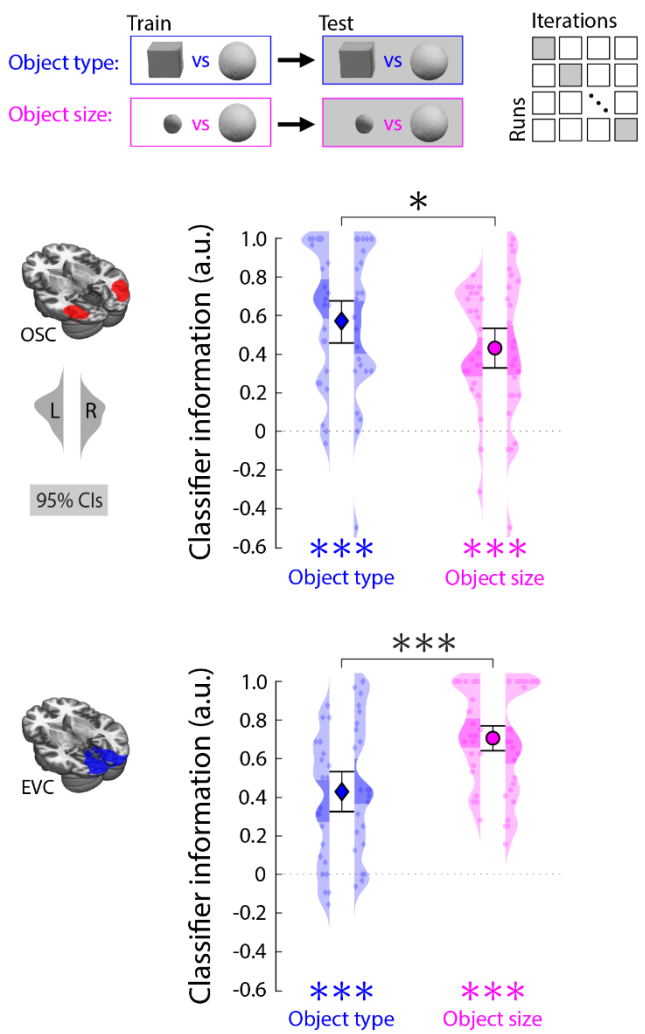
Online supplements

Online Supplement OS1: Decoding of object type and object size in training runs.

In this study we sought to characterize how a search target is represented in neural activity during search preparation. In the main manuscript, we therefore reported the results of cross-classification analyses, showing that classifiers trained on visually evoked activity (i.e., viewing melons versus boxes in the training runs), successfully generalized to activity evoked during search preparation (i.e., searching for melons versus boxes in the search task). A prerequisite for such cross-generalization is that the visually presented objects evoke distinguishable patterns of visually evoked activity. In this supplement, we tested whether our regions of interest (ROIs) meet these requirements, by quantifying the amount of information about the visual objects within activity patterns evoked during the training runs (Supplemental Figure S1). Performing a leave-one-run-out cross-validation approach, we show that classifiers could distinguish between activity patterns evoked by viewing a melon versus a box, in both OSC and EVC (both $p < .0005$). This establishes that our ROIs comprised information about the type of object that was visually presented.

In the subsequent Results section of the main manuscript, we compared this object-based cross-classification performance between two training regimes. Specifically, we compared cross-classification of the target object (melons versus box) between a classifier that was trained on visual objects of a size that matched the current search distance (i.e., small objects for distant search, and large objects for nearby search) and a classifier that was trained on size-mismatching visual objects (vice versa). In order for these two training regimes to alter performance, it is required that visually presented objects of different sizes evoke distinguishable patterns of visually evoked activity. Using the same approach as above, we show that classifiers could distinguish between activity patterns evoked by viewing small objects versus large objects, in both OSC and EVC (both $p < .0005$). This establishes that our ROIs comprised information about the type of object that was visually presented. Note that OSC activity comprised more information about object type than object size ($p = .029$), whereas the opposite was true for EVC ($p < .0005$), thus reflecting the differential specialization of these two brain regions.

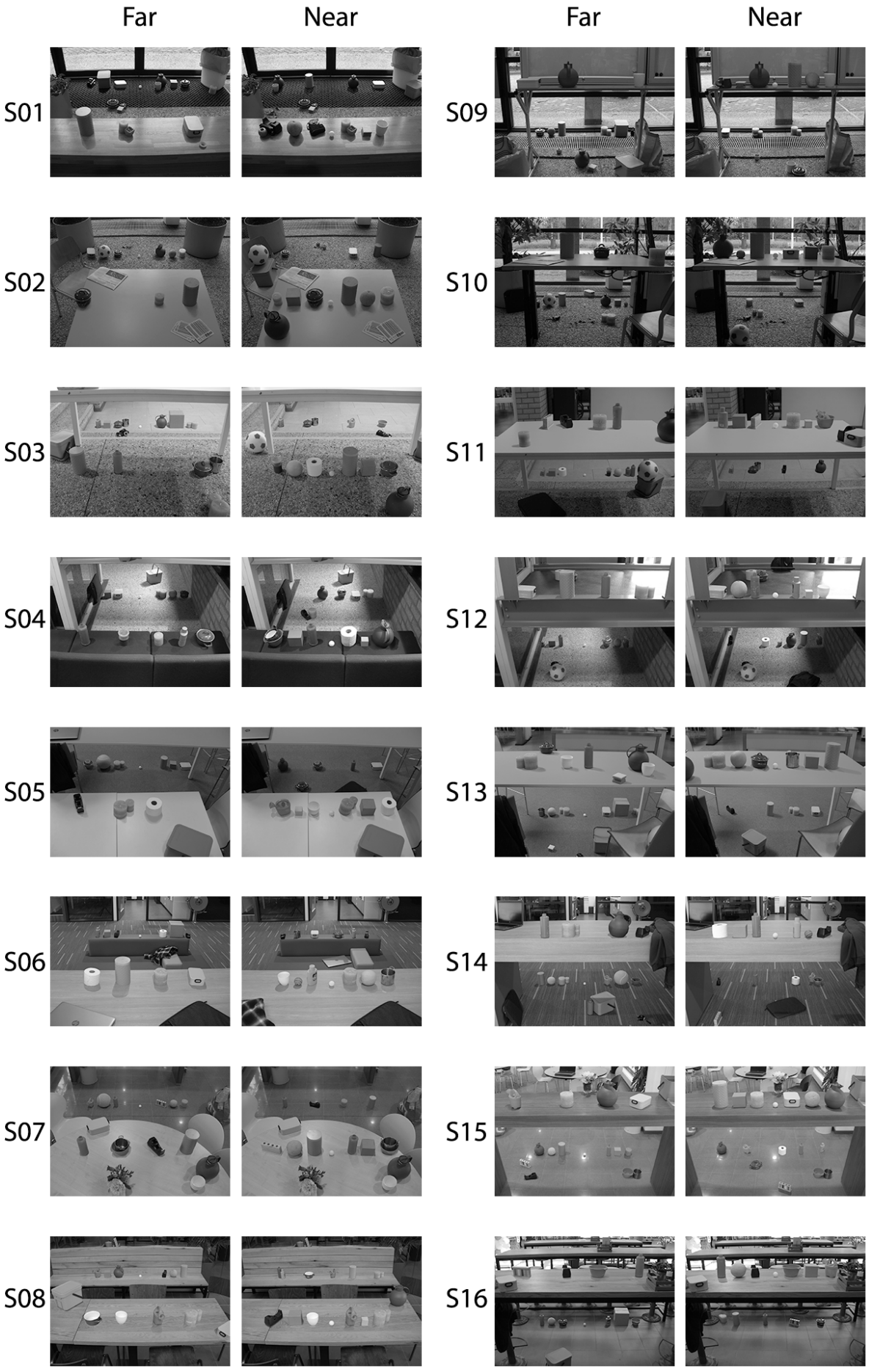
In sum, these analyses show that both of our ROIs comprised information about the size and type of our target objects, at least when visually presented. Both ROIs therefore met the functional requirements to allow for (1) maintaining object-specific representations during search preparation, and (2) rescaling these representations to account for viewing distance during search.



Online Supplement Figure OS1. Information about object type and object size in visually evoked activity. We tested whether patterns of visually evoked activity in OSC and EVC distinguished between visual object types (i.e., images of melons versus boxes) and visual object sizes (i.e., small versus large images of melons and boxes) in the training runs. Multivariate classification was achieved using a linear support vector machine (libsvm) on run-based beta-maps following a leave-one-run-out cross-validation procedure (akin to the analysis approach reported in the Results section of the main manuscript). Small colored dots represent classifier information (derived from distance-to-bound) for individual participants, obtained separately from the left and right hemispheres (displayed within the left and right kernel-density plots, respectively). The central markers reflect the population mean, averaged across hemispheres. Error bars around the central markers, and shaded areas within the kernel-density plots represent the bootstrapped 95% confidence intervals of the mean. *p < .05, **p < .005, ***p < .0005.

Online Supplement OS2: Example stimuli from each of the 16 scene families.

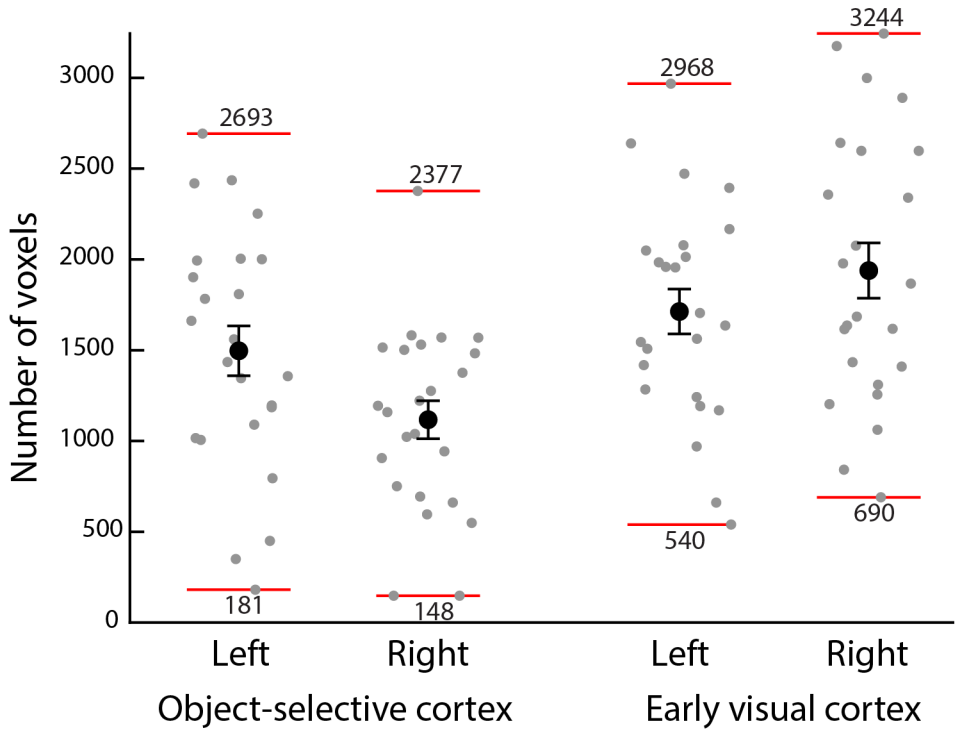
A total of 512 unique search arrays were created, based on 16 different scene families (see Supplemental Figure S4). This allowed us to present participants each unique search array only once over the course of the entire two-session experiment. Note that the results of the fMRI analyses reported throughout the manuscript were based on a subset of trials (i.e., 50%) in which *no* array of objects was presented in these 16 scenes. Thus, preparatory activity was measured while observers stared at these scenes when they were devoid of objects.


Far
Near
Far
Near
S01
S02
S03
S04
S05
S06
S07
S08
S09
S10
S11
S12
S13
S14
S15
S16

Online Supplement Figure S2. Example stimuli of the search task, in which participants were cued to search for a melon or a box. The stimulus set was composed of 16 different scene families (labeled S1-S16) with a distinct scene background. For each scene, we depict an example of a search trial in which participants were cued to search on the far plane (left side of each column) and on the near plane (right side of each column). In half of the scenes (S1-S8, on the left) the 'near' location was below the 'far' location, and in the other half of the scenes (S9-S16) the 'far' location was below the 'near' location. This allowed us to decouple the search distance from the location cue (which indicated the top or bottom location). For each of these 16 scene families and two search distances, we created 16 different arrays of objects (totaling to 512 unique images).

Online Supplement OS3: Number of voxels in regions of interest per participant.

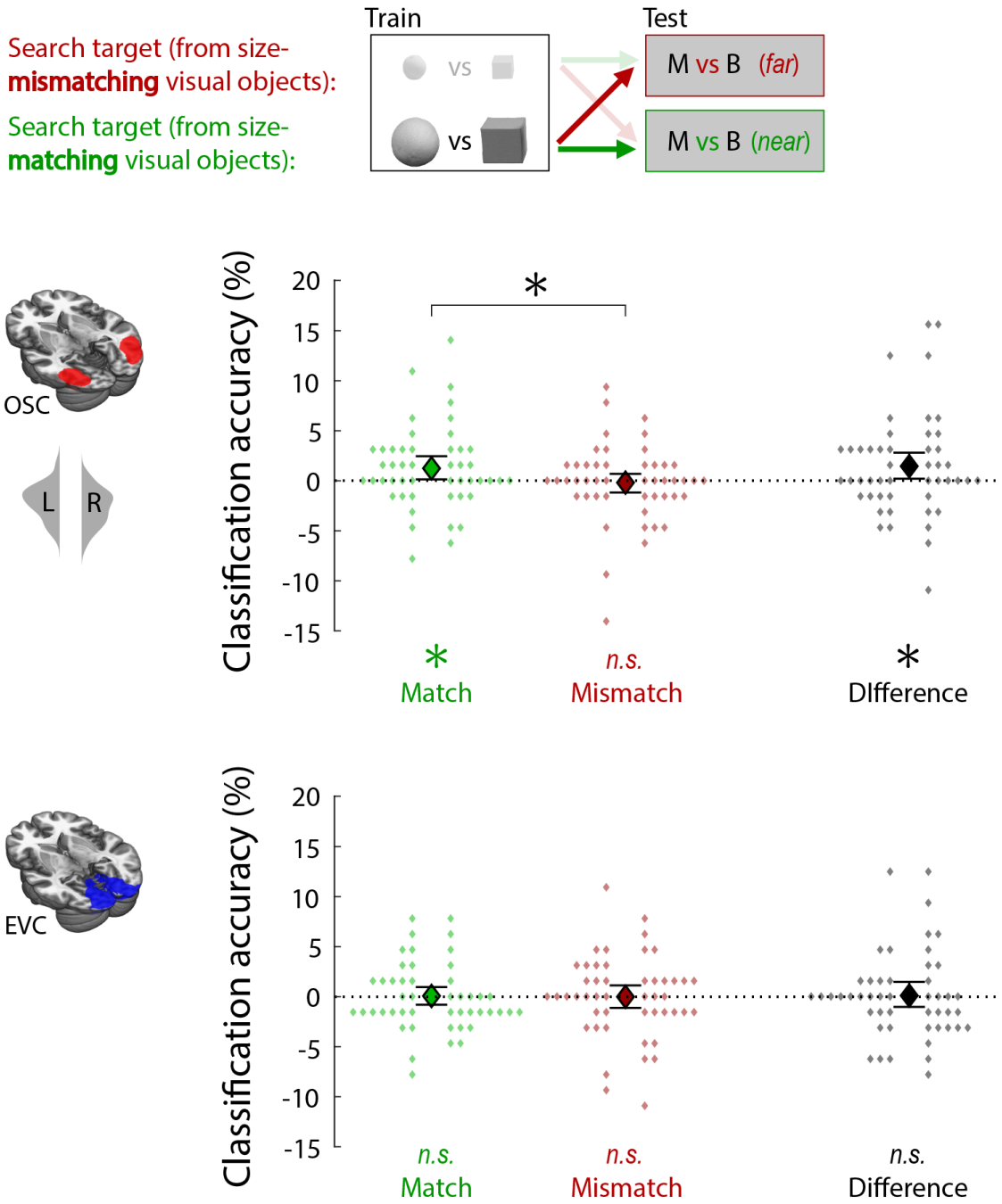
Regions of interest (ROI) were defined per participant, by intersecting group-level masks with voxels that showed reliable ($p_{uncorrected} < .05$) object-selective (for OSC) and visually-driven (for EVC) responses at the individual level. As a result, individual participants' ROIs comprised different numbers of voxels. In Supplemental Figure S5, we show the number of object-selective and visually-responsive voxels within the group-level masks, separated by hemisphere. As can be seen, although the number of voxels varied greatly per participant and between hemispheres, each participant had a substantial number of included voxels (at least 148) in each hemisphere of each ROI. The OSC ROI had an average size of 1497 voxels (SD = 672) in the left hemisphere, and of 1173 voxels (SD = 512) in the right hemisphere. The EVC ROI had had an average size of 1713 voxels (SD = 604) in the left hemisphere, and of 1939 voxels (SD = 747) in the right hemisphere.



Online Supplement Figure OS3. Number of voxels included in both ROIs (object-selective cortex and early visual cortex), separated by hemisphere (left and right) for each individual participant (gray dots). For object-selective cortex, this reflects voxels that responded more to objects than scrambled objects in the functional localizer runs, within a population-level functionally-defined object-selective mask (retrieved from Julian, Fedorenko, Webster, & Kanwisher, 2012). For early visual cortex, this reflects voxels that responded to our target stimuli in the training runs (as compared to the implicit baseline), within Brodmann's Areas BA17 and BA18. The red lines and corresponding numbers indicate the minimum and maximum number of voxels for that hemisphere and ROI. The black markers indicate the population mean, and the error bars reflect the standard error of the mean.

Online Supplement OS4: Main results in classification accuracy

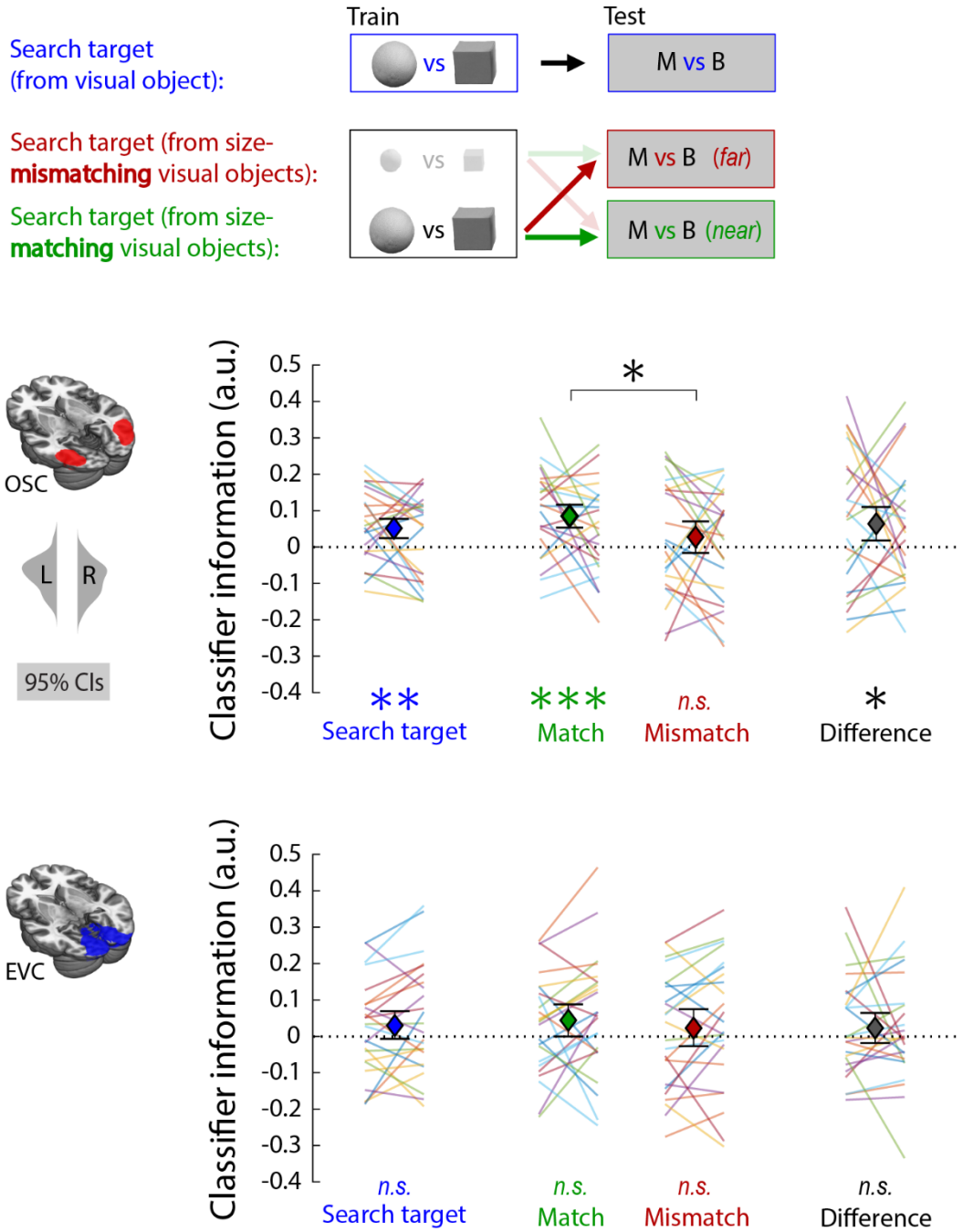
In the main manuscript file we report our results in a classifier information metric derived from the distance-to-bound values of the classifier outcome. This is known to be more sensitive than metrics based on binary classification outcome. In our case in particular, binary classification metrics are less sensitive due to an inherent classification bias: classifiers trained on activity evoked by viewing isolated objects (in the training runs), were biased to classify activity evoked during search preparation (in the search task runs, while participants viewed an empty scene) as "box" on a large proportion of the trials; up to 90% for individual participants. Given such a strong bias, classification accuracies (1) have a theoretical ceiling performance that is much lower than 100% (even given data without noise), and (2) are effectively based on much fewer data points (i.e., only those test examples that are sufficiently "melon-like" to overcome the bias), and are therefore inherently more noisy. Our current metric of classifier information, instead, reflects how much more box-like the classifier outcome was when classifying boxes compared to when classifying melons, thus (1) implicitly removing any such bias, and (2) effectively using all datapoints. Nonetheless, we share the results of our analysis using the more traditional metric of binary classification accuracy metric, to show that (statistically speaking) our results hold when using this suboptimal method.



Online Supplement Figure OS4. To corroborate the key cross-classification results reported in the main manuscript (Figure 3), we replicated our analyses using the more traditional metric of binary classification accuracy. Specifically, we tested whether a classifier trained to distinguish between activity evoked by viewing isolated images of melons versus boxes (training runs) could distinguish search for melons versus boxes in the search task runs. This was done in two conditions: we tested whether this melon versus box cross-classification was better when training on the appropriate object size (in green) compared to the inappropriate object size (in red) given the viewing distance during search (near of far). The within-subject difference between these two conditions is depicted in black. Our key findings, as reported in the main manuscript were replicated using this approach. Small colored dots represent an individual subject's mean classification accuracy, separately obtained from the left and right hemispheres (displayed to the left and right of the central marker, respectively). The central markers reflect the population mean, averaged across hemispheres. Error bars around the central markers represent the bootstrapped 95% confidence intervals of the mean (2000 permutations). *p < .05.

Online Supplement OS5: Main results in classification accuracy

Each classification analysis in the main manuscript was performed for each observer in the left and right hemispheres separately (and was later averaged to perform statistical testing). Here, we display the same results as in Figure 2B and 3 of the main manuscript, but connected the results obtained from the left and right hemispheres of the same participants. This shows the consistency between hemispheres of the observed effects. In general, interhemispheric consistency was weak in object-selective cortex and moderate in early visual cortex, but was positively correlated in both regions of interest.



Online Supplement Figure OS5. Consistency between hemispheres of individual participants depicted for the key findings from the main manuscript (Figures 2B and 3). Colored lines depict the individual subjects' mean classifier information for the left and right hemispheres (displayed to the left and right of the central marker, respectively). The central markers reflect the population mean, averaged across hemispheres. Error bars around the central markers represent the bootstrapped 95% confidence intervals of the mean (2000 permutations). *p < .05, **p < .005, ***p < .0005.