

Position and Identity Information Available in fMRI Patterns of Activity in Human Visual Cortex

 Zvi N. Roth^{1,2} and Ehud Zohary^{1,2}

¹Department of Neurobiology, Life Sciences Institute, and ²The Edmond and Lily Safra Center for Brain Sciences, Hebrew University, 91904 Jerusalem, Israel

Parietal cortex is often implicated in visual processing of actions. Action understanding is essentially abstract, specific to the type or goal of action, but greatly independent of variations in the perceived position of the action. If certain parietal regions are involved in action understanding, then we expect them to show these generalization and selectivity properties. However, additional functions of parietal cortex, such as self-action control, may impose other demands by requiring an accurate representation of the location of graspable objects. Therefore, the dimensions along which responses are modulated may indicate the functional role of specific parietal regions. Here, we studied the degree of position invariance and hand/object specificity during viewing of tool-grasping actions. To that end, we characterize the information available about location, hand, and tool identity in the patterns of fMRI activation in various cortical areas: early visual cortex, posterior intraparietal sulcus, anterior superior parietal lobule, and the ventral object-specific lateral occipital complex. Our results suggest a gradient within the human dorsal stream: along the posterior–anterior axis, position information is gradually lost, whereas hand and tool identity information is enhanced. This may reflect a gradual transformation of visual input from an initial retinotopic representation in early visual areas to an abstract, position-invariant representation of viewed action in anterior parietal cortex.

Key words: dorsal stream; fMRI; intraparietal sulcus; MVPA

Significance Statement

Since the seminal study of Goodale and Milner (1992), there is general agreement that visual processing is largely divided between a ventral and dorsal stream specializing in object recognition and vision for action, respectively. Here, we address the specific representation of viewed actions. Specifically, we study the degree of position invariance and hand/object manipulation specificity in the human visual pathways, characterizing the information available in patterns of fMRI activation during viewing of object-grasping videos, which appeared in different retinal locations. We find converging evidence for a gradient within the dorsal stream: along the posterior–anterior axis, position information is gradually lost, whereas hand and action identity information is enhanced, leading to an abstract, position-invariant representation of viewed action in the anterior parietal cortex.

Introduction

The functional role of a sensory region in the brain is typically assessed by identifying the stimulus dimensions for which a change in specific parameters modulates the neuronal response. This “tuning” portrays the degree of selectivity to a specific fea-

ture (e.g., stimulus size or color). Equally important is the complementary aspect, a lack of sensitivity to changes in a specific parameter, which indicates the degree of invariance to that parameter. The degree of sensitivity to a specific dimension can be studied both at the level of a single “unit” (be it a neuron or a voxel) or at the population (“vector”) level. A generalizing unit responds to different stimuli varying along a specific parameter at a constant activity level, whereas a selective unit shows clear variation in its response. Analogously, a generalizing population responds with similar patterns of activity (across units) to different stimuli, whereas selectivity entails a consistently different pattern of activity for the various stimuli.

The selectivity of single units and their population response vector to stimuli in the ventral visual pathway has been tested systematically using univariate and multivariate approaches, re-

Received Feb. 24, 2015; revised July 2, 2015; accepted July 10, 2015.

Author contributions: Z.N.R. and E.Z. designed research; Z.N.R. performed research; Z.N.R. analyzed data; Z.N.R. and E.Z. wrote the paper.

This study was funded by an Edmond & Lily Safra Center for Brain Sciences (ELSC)/Ecole polytechnique fédérale de Lausanne (EPFL) Research Grant to E.Z. We thank Tanya Orlov for helpful discussions and Yuval Porat for assistance with collection of eye movement data.

Correspondence should be addressed to either Zvi N. Roth or Ehud Zohary, Department of Neurobiology, Life Sciences Institute, Hebrew University, Edmond J. Safra Campus, 91904 Jerusalem, Israel, E-mail: zvi.roth@mail.huji.ac.il or udiz@mail.huji.ac.il.

DOI:10.1523/JNEUROSCI.0752-15.2015

Copyright © 2015 the authors 0270-6474/15/3511559-15\$15.00/0

spectively. For example, Rust and Dicarlo (2010) found that the response vector, based on a population of hundreds of neurons, showed increased selectivity to image identity along the cortical hierarchy (i.e., when comparing V4 and IT population responses). Conversely, the population vector displayed increased invariance to changes in the image retinal size and its position in the visual field. Similarly, an analysis of the multivoxel patterns of activity evoked by various stimuli has been used to assess category selectivity in the human ventral pathway (Haxby et al., 2001; Kriegeskorte et al., 2008). For example, recent work using multivoxel pattern analysis (MVPA) revealed that higher-order visual regions in ventral occipitotemporal cortex display clear category and exemplar selectivity while manifesting relative invariance to changes in the retinal position in which the stimuli appeared (Cichy et al., 2013).

The degree of selectivity and generalization in the dorsal pathway has generally been neglected. Valyear et al. (2006) found, using fMR adaptation, that responses to still object images in the occipitoparietal junction were selective to object orientation but were invariant to object identity, whereas the opposite pattern of selectivity was found in the lateral occipital complex (LOC). However, it is still unclear exactly which stimulus properties determine visual responses in parietal regions. It is also currently unknown how sensitive parietal regions are to the location of viewed actions in the visual field. Identity and location sensitivity should reflect the different functional roles of various parietal regions. For example, a brain area involved in planning or performing actions is expected to contain information regarding an object's location because this information is crucial for performing an action with that object. If, however, a region is involved in understanding actions made by others rather than planning one's own actions, then accurate localization might not be of importance. In that case, the region's activity is more likely to carry information regarding the object's identity that is essential for action understanding. Therefore, revealing the generalization and selectivity to location and tool identity in parietal regions should help us to elucidate these areas' potential functional roles. To that end, we characterized the information available in the patterns of activation when viewing object-grasping actions in various cortical regions.

Materials and Methods

Subjects. Fifteen healthy right-handed subjects (four females) gave their informed consent to participate in the fMRI study, which was approved by the Helsinki Ethics Committee of Hadassah Hospital, Jerusalem, Israel. One subject was excluded from further analysis due to excessive head movement (>2 mm) during one of the scans. Therefore, the data from 14 subjects were used in this study.

Stimuli. Each of six 1800 ms video clips depicting a right hand grasping and using a tool (hammer, screwdriver, stapler, corkscrew, garlic press, or knife), was downsampled to 140×140 pixels. The original clips featured a right hand and, by creating mirror-image clips, six left-hand clips were generated, resulting in a total of 12 different clips (Fig. 1a, bottom).

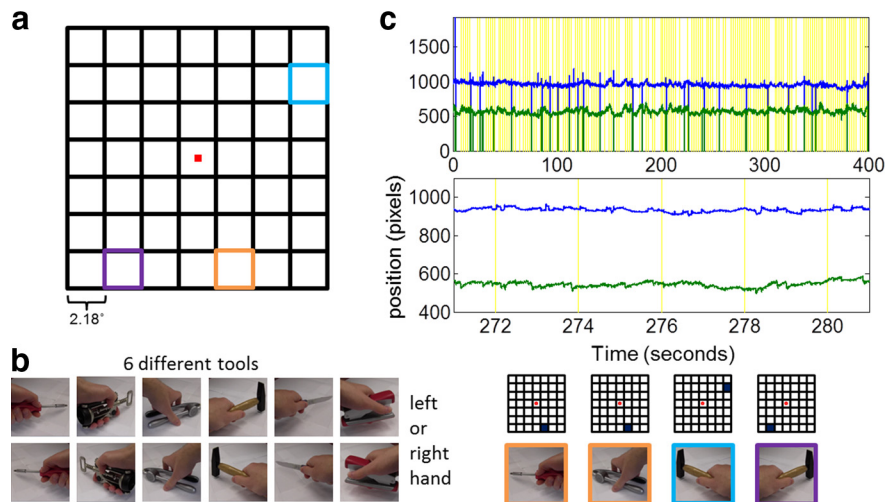


Figure 1. Experimental design and eye tracking. **a**, In each trial, a video clip appeared in one of 49 possible locations while the subject kept fixation on the central red dot during the entire run. The grid, indicating the 49 locations, was not shown during the experiment. **b**, Characteristic still images from each of the 12 clips used. The top six clips show a left hand grasping a tool and the bottom ones show the identical clips after a mirror image inversion, i.e., the right hand grasping the same tools. **c**, Eye-tracking results for one representative subject. Top, Eye position traces for the first run. Horizontal position is shown in blue and vertical position in green. Vertical yellow lines represent the beginning of each trial (excluding null trials). Middle, Close-up on four trials. Below the graph, grids schematically indicate the location of the stimulus during each trial, still images indicate the clip presented during each trial, and each image is outlined by the color corresponding to the clip location in **a**. Although the stimulus position changes across trials, eye position remains successfully fixated.

Experimental design and tasks. Each subject completed 8 runs of the main experiment and a localizer run across 2 scanning days. During each run of the main experiment, subjects fixated 2 on a central red square while viewing video clips at various locations on the screen. Subjects were instructed to covertly name both the hand (left or right) and the tool in each clip without moving their eyes. Each session began with a training period during which subjects first fixated on the clips until they felt acquainted with them and then practiced keeping central fixation while paying attention to the clips and covertly naming them. This training continued until the subjects reported feeling comfortable with the task and only then did the scanning commence. During each event-related run of the main experiment, three clips appeared in each of the 49 possible locations (Fig. 1a, left): eight of the 12 clips appeared 12 times (in 12 different locations) and the other four clips appeared 13 times (in 13 different locations). Across all eight runs, each clip was presented 98 times, twice in each location, and each location hosted 24 clip presentations with three presentations per run. Each trial lasted 2000 ms [1800 ms clip duration with 200 ms interstimulus interval (ISI), i.e., 1 TR]. In addition to 147 clip trials, each run included 49 randomly interspersed null trials during which no clip was presented and subjects maintained fixation on the central red square. Four additional null trials at the beginning and at the end of the run brought the total run duration to 6:48 min (204 volumes).

The localizer scan was composed of six blocks of hand, face, animal, tool, and phase-scrambled images. Each block consisted of 32 images, presented for 450 ms each with 50 ms ISI. In each block, zero to two images were shown consecutively and subjects indicated such repetitions by button press (i.e., a one-back task). Four initial null trials and four final null trials brought the run duration to a total of 8:16 min (248 volumes).

Eye tracking. During most of the runs, eye movements were recorded and monitored online via a video-based, infrared eye tracker (Eye Link1000; SR Research) with a sampling rate of 1000 Hz. This enabled us to make sure that subjects were fixating the center and were awake and alert. Unfortunately, due to technical problems, the signal was often noisy and unstable. Nevertheless, visual examination of the eye position time courses in the times in which the eye position signal was reliable revealed that subjects generally complied with instructions and kept central fixation (see example time course in Fig. 1c). To confirm that the stimulus location did not heavily skew the eye position, for each subject,

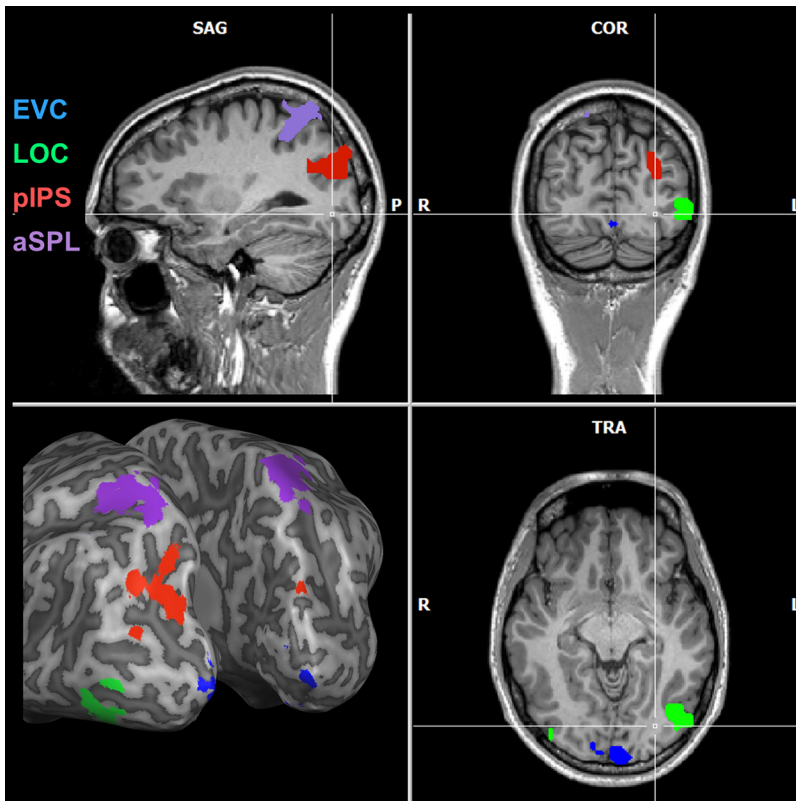


Figure 2. ROIs consisting of the EVC (blue), LOC (green), pIPS (red), and aSPL (purple). Top left, top right, and bottom right panels present the sagittal, coronal, and transverse views, respectively, in radiological convention (right hemisphere is to the left of the image) for one representative subject. Bottom left, Same ROIs (depicted in their corresponding colors) shown on an inflated version of the same subject's brain from a posterior, slightly right-rotated viewpoint.

we averaged the gaze location across all trials during which the stimulus was presented in each visual field quadrant. The differences between quadrants were small, both relative to the stimulus size and to the SDs, and we did not find a consistent trend across subjects.

MRI scanning parameters. The blood oxygenation level dependent (BOLD) fMRI measurements were obtained using a 3-T Magnetom Skyra Siemens scanner and a 32-channel head coil. The functional MRI protocols were based on multislice gradient echoplanar imaging and obtained under the following parameters: TR = 2 s, TE = 30 ms, flip angle = 90 degrees, imaging matrix = 64×64 , FOV = 192 mm; 37 slices with 3 mm slice thickness and 15% gap (0.45 mm) were oriented in an oblique position covering the whole brain, with functional voxels of $3 \times 3 \times 3$ mm. In addition, high-resolution T1-weighted magnetization-prepared rapid acquisition gradient-echo images were acquired ($1 \times 1 \times 1$ mm resolution). All scans used GRAPPA parallel imaging (acceleration factor = 2).

Data processing. Data analysis was conducted using the BrainVoyager QX software package (Brain Innovation) and in-house analysis tools developed in MATLAB (The MathWorks). Preprocessing of functional scans included 3D motion correction, slice scan time correction, and removal of low frequencies (linear trend removal and high-pass filtering). The anatomical and functional images were transformed to the Talairach coordinate system using trilinear interpolation. The cortical surface was reconstructed from the high-resolution anatomical images using standard procedures implemented with BrainVoyager software.

Voxel time courses were generated using BrainVoyager and analyzed using MATLAB custom-made software. Specifically, we first transformed each voxel's time course to obtain a z-score value by subtracting the mean activation and dividing by the SD of the BOLD response across the whole run. Next, we used a standard general linear model (GLM) analysis with a regressor for each condition, assuming the standard (two gamma) hemodynamic response function (Friston et al., 1998). This

resulted in one activation parameter (β -value) per condition for each voxel. In each of the identity and location analyses (detailed below), we subdivided the experimental conditions into different conditions, thereby obtaining a different number of regressors or predictors. To each regressor, we assigned a specific β -value based on its activation level. We then transformed the β -values into t -values by subtracting each voxel's mean β -value (across all conditions) and dividing by each β -value's SD.

ROI selection. Using a functional localizer we defined four ROIs: early visual cortex (EVC), LOC, posterior intraparietal sulcus (pIPS) and anterior superior parietal lobule (aSPL) (Fig. 2). To obtain two separable ROIs in the parietal per subject, we used the contrast hands > scrambled and, starting from FDR < 0.001, raised the threshold separately for each hemisphere until a cluster of active voxels in the anterior portion of the parietal cortex was separated from the pIPS. We then raised the threshold further (if necessary) until a cluster of pIPS voxels was separated from occipital areas and thus defined pIPS. Next, we contrasted tools > scrambled and, starting from FDR < 0.001, raised the threshold until a ventral-occipital cluster separated from the parietal region, and thus selected LOC. Finally, we used the fact that early visual areas are sensitive to local contrast that is enhanced in the scrambled images. Therefore, the opposite contrast (scrambled > tools) and the same threshold as for LOC was used to select a cluster of voxels in the posterior occipital cortex, defined as EVC.

When we performed the analyses separately on the right and left hemisphere ROIs, the results across regions remained similar, although there were small differences between the hemispheres. These differences, however, were usually consistent across the ROIs (e.g., all left hemisphere ROIs showed slightly higher tool identity classification results than all right hemisphere ROIs). Therefore, we report only results from analyses on bilaterally defined ROIs.

GLMs. To assess the image-specific characteristics of the studied ROIs, we performed four different analyses, each with a different set of predictors. This resulted in four different sets of coefficients (β -values) that were transformed into t -values. For each voxel, the regressors represented the location (49 coefficients), the hand identity (two coefficients), the tool identity (six coefficients), or the clip identity (i.e., combined hand and tool; 12 coefficients).

MVPA. For each of the four analyses, we used two methods of MVPA: correlation analysis and support vector machine (SVM) classification. Although these two methods probe slightly different aspects of the multivoxel patterns, we expected their results to largely correspond to each other (Weber et al., 2009; Golomb and Kanwisher, 2012).

For the correlation analysis, the eight runs were split into two ("even" and "odd") groups of four runs each and, for each group, the runs' time courses were concatenated. We then performed the GLM regression and normalized each voxel's t -values by subtracting the mean across conditions for that voxel (Haxby et al., 2001; Garrido et al., 2013; Roth and Zohary, 2014; repeating all correlation method analyses without this normalization yielded similar results). This resulted in two sets of t -values, one for each half of the runs (Fig. 3a). Correlating all of the patterns of activation (one pattern for each condition) in half of the runs with all of the patterns in the other half results in a correlation matrix of size $p \times p$, where p is the number of conditions. Subtracting the mean overall correlations between different conditions from the mean overall correlations between identical conditions gives a measure of the stimulus information carried by the patterns. To minimize the effects of random

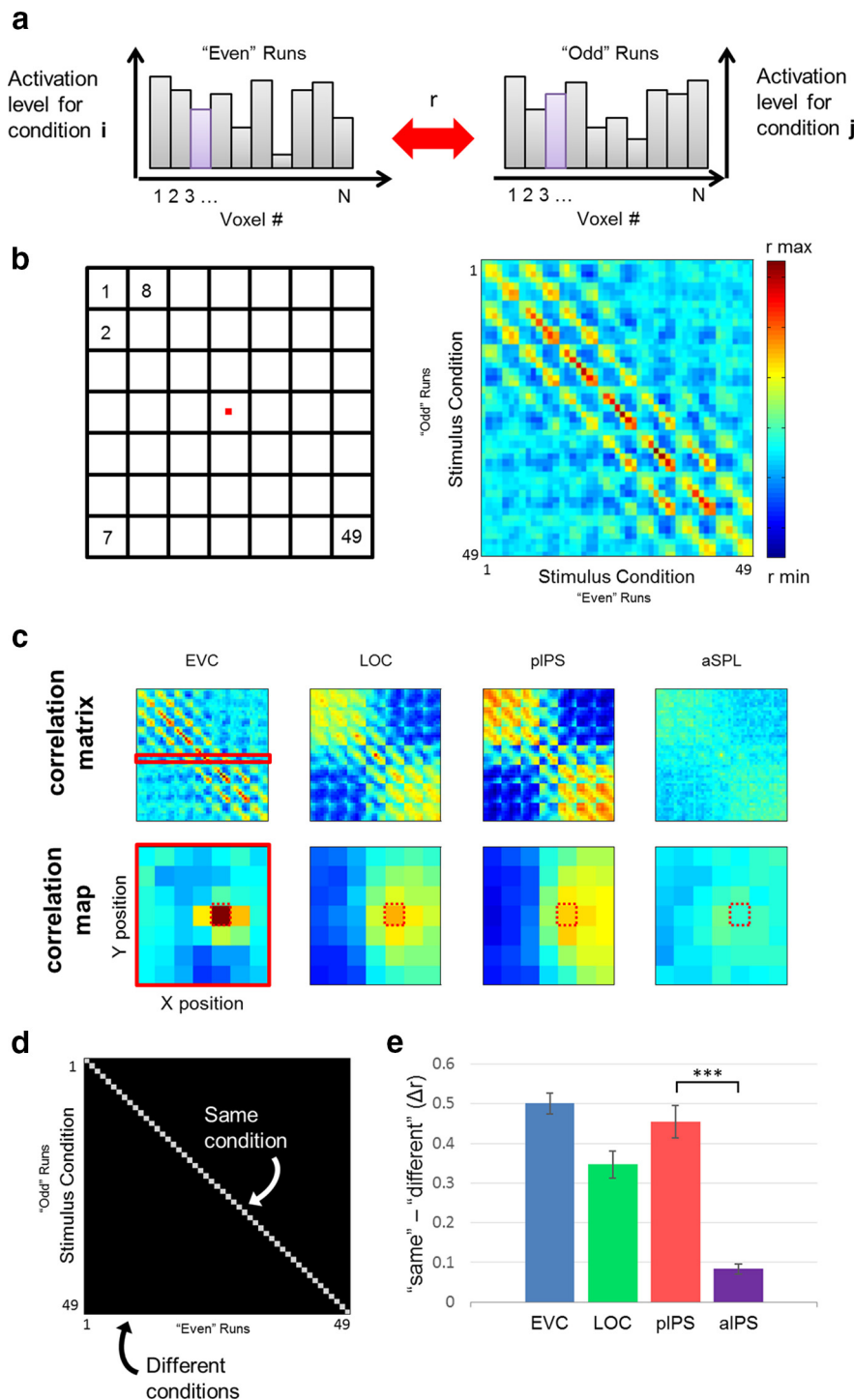


Figure 3. Location analysis: correlation method. *a*, In MVPA, each condition is represented by a vector of activation levels across voxels within a specific ROI. For the correlation method, we computed the correlation between the vectors of every condition in one half of the runs (e.g., “even” runs) with the vectors of every condition in the other half (e.g., “odd” runs). *b*, Location correlation matrix. Left, The 49 locations are numbered in the order in which they appear in the correlation matrices. Right, The correlation matrix of EVC, averaged across all subjects’ ($N = 14$) individual matrices. Note that the correlations are greater along the main diagonal (same position) and the parallel lines that reflect the nearest neighbors in terms of stimulus positions. *c*, Top, Location correlation matrices for each ROI, averaged across all subjects’ individual matrices. Bottom, Location correlation maps for each ROI. The correlation map is created by reordering a single row (#32 in this example) or column from the correlation matrix into a smaller 7×7 matrix that reflects the correlation between the pattern evoked by a stimulus in each position, with the activity evoked by a stimulus in position #32. It explicitly depicts the dependence of the correlation on the spatial position of the stimuli. The correlation becomes less dependent on the exact position of the stimuli from EVC to aSPL, reflecting the increasing size of the receptive fields in parietal cortex. *d*, We measured the information of an ROI by averaging the diagonal (correlations between the patterns of activity evoked by stimuli at the same location) and subtracting the average of all off-diagonal values (correlations between the patterns of activity evoked by stimuli at different locations). *e*, Results of the location analysis using the correlation

differences between the two groups of runs, we repeated this analysis 12 times, each time splitting the runs into different groups. The correlation values were then averaged across splits. To test for statistically significant differences between same-condition correlations and different-condition correlations, mean correlation values (across 12 splits and across conditions) for each subject were transformed to Fisher z -scores, and t tests were performed on the difference in z -scores between different conditions and identical conditions (ANOVA revealed a significant effect of ROI for all analyses and both MVPA methods). To test for significant differences between regions, paired t tests were performed on the Fisher z -scores of three pairs of regions: LOC and pIPS, LOC and aSPL, and pIPS and aSPL. To avoid clutter, in all graphs, we present only the results for the pIPS versus aSPL t test comparison. These p -values were corrected for three comparisons.

For the classification analysis, the GLM regression was performed separately for each run and the resulting t -values were normalized to the range of $[0, 1]$. Binary classification SVMs were trained on the data from seven runs and tested on the final run in a leave-one-run-out cross-validation method using LIBSVM (Chang and Lin, 2011). Whereas the hand identity analysis involved an obvious classification of each trial as either a left-hand or right-hand clip, the binary classification for the other analyses was less obvious because there are multiple possible ways to divide the conditions into two equal-sized groups. We therefore used several classifications for each analysis, eventually averaging the results. For the tool and clip identity analyses, we used all possible classifications (10 and 462, respectively), whereas for the location analysis, we used 128 random classifications (similar results were obtained using 256 and 512 random classifications). For the location classification analysis, to have an even number of locations, we ignored the central location, leaving 48 locations to be classified. In addition to the random splits, we trained classifiers on more orderly splits of the locations, which were nonrandomly chosen to demand varying levels of fine-scale discrimination between locations (Fig. 4*d*). We used t tests to determine whether accuracies were significantly greater than chance and paired t tests to test for significant differences between regions’ accuracies (as for the correlation method: LOC and pIPS, LOC and aSPL, and pIPS and aSPL).

Generalization across hemifields. If patterns of activation contain stimulus information that is location independent, then patterns corresponding to certain stimuli occurring in one-half of the visual field should be highly similar to patterns corresponding to the same stimuli in the other half of the visual field. To determine whether such location invariance is char-

method. Scale bar values and error bar values here and in subsequent figures represent the mean across subjects \pm SEM. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns: $p > 0.05$, corrected for three comparisons.

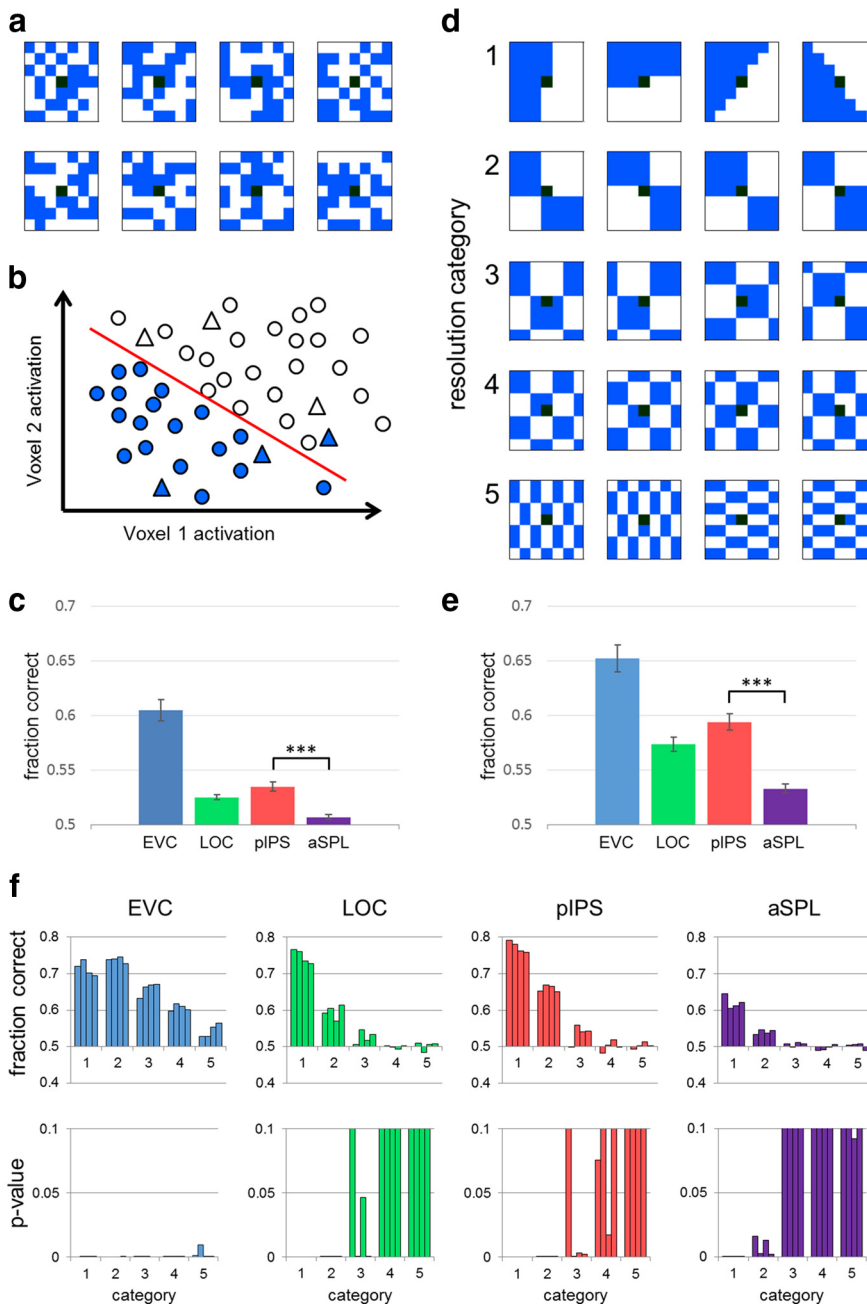


Figure 4. Location analysis: classification method. *a*, Eight examples of random location classifications. The 48 locations (the center location is left out of this analysis) are randomly divided into two equal-sized groups, which are defined as the two classes to be classified by the SVM. *b*, Classification method. Each location is represented by a vector of voxel activation levels; that is, a vector of length N , where N equals the number of voxels in the ROI. In the schematic example, the ROI consists of only 2 voxels, and each location is represented by a vector of length 2. The circles represent locations from the training data from both classes (blue and white). The SVM finds a straight line (or an $N-1$ hyperplane in the general case of N voxels) that separates the two classes and is then tested on locations from the testing data, represented by triangles. In this example, the classifier is correct in 5/6 cases of the testing data (the line correctly classifies 5 of the triangles), resulting in a score of 0.83. *c*, Results of the random classifications analysis. *d*, All 20 nonrandom classification, grouped into 5 resolution levels, from the category demanding the lowest resolution (#1) to the highest (#5). *e*, Results of the nonrandom classifications analysis, averaged across subjects. *f*, Nonrandom classification results grouped by resolution category. Top, Classification accuracy for each ROI and resolution category. Bottom, p -values of t tests for each ROI and resolution category. Patterns of activity in aSPL contain enough location information to correctly classify the first two resolution categories.

acteristic of our ROIs, we divided the 49 location into three groups: left visual field (LVF), vertical meridian, and right visual field (RVF). We ignored all trials located in the vertical meridian and performed MVPA across the two hemifield. For the correlation analysis, we correlated patterns from LVF with patterns from RVF and vice versa (see Fig. 7*a*),

whereas for the classification analysis, we trained SVMs to differentiate between conditions (e.g., right hand vs left hand) in the LVF and tested them on the RVF (and vice versa).

Generalization across hands. We performed a similar generalization analysis to test for tool identity information that is independent of hand identity. To that end, we split each voxel’s 12 t -values (corresponding to each of the video clips), which were computed for the clip identity analysis according to the acting hand in the clip (resulting in six left-hand t -values and six right-hand t -values). For the correlation method, we correlated the multivoxel patterns evoked by right-hand clips with the left-hand patterns and vice versa (see Fig. 8*a*) and then subtracted the mean off-diagonal values (“different”) from the mean on-diagonal values (“same”). For the classification method, we trained SVMs to differentiate between the six clips of tools being manipulated by the left hand and tested them on the six right-hand clips for every one of the 10 classifications. The same procedure was done for the opposite direction (training with right hand, testing with the left-hand clips). The results were averaged across the various classifications and trained hand.

Searchlight analysis. As a complementary analysis and to verify that our ROIs captured the regions with relevant information, we performed a searchlight analysis (Kriegeskorte et al., 2006), performing the correlation analysis using cubes of voxels instead of the predefined ROIs. Specifically, we used cubes measuring $5 \times 5 \times 5$ and $7 \times 7 \times 7$ voxels ($15 \times 15 \times 15$ and $21 \times 21 \times 21$ mm³, respectively) and used every possible cube of contiguous voxels. As with the ROI-based correlation analysis, we split the runs in 12 different ways and created a correlation matrix for each split. We then averaged the matrices across splits and subtracted the Fisher-transformed mean correlation between different conditions (off-diagonal) from the Fisher-transformed mean correlation between identical conditions (the diagonal of the correlation matrix) for each subject. Performing t tests on the difference values left us with a t -value for each searchlight center voxel. Results for both sizes of the searchlight cubes were similar, so we present only results for the smaller cube.

Results

We studied the patterns of activity evoked by various dynamic visual stimuli in select regions of the dorsal and ventral pathways: EVC, parietal areas pIPS and aSPL, and the ventral object-specific constellation LOC. Our participants viewed video clips depicting a hand grasping various tools and using them in their characteristic way at 49 different screen (and retinotopic) locations (Fig. 1*a*). This allowed us to obtain a comprehensive measure of the available information about the visual stimulus structure and its location in the visual field. We used two forms of MVPA, correlation analysis and pattern classification

(Fig. 1*a*). This allowed us to obtain a comprehensive measure of the available information about the visual stimulus structure and its location in the visual field. We used two forms of MVPA, correlation analysis and pattern classification

(with SVMs), to assess the information available for discrimination either between stimulus locations or between different object/action attributes (hand or tool identity or the specific combination of the two).

Location analysis

As a first step, we wanted to determine the degree of sensitivity of visual cortical regions to the position of visual stimuli on the screen. Accurate localization (e.g., where is the tool) is obviously crucial for performing actions, but is not necessarily relevant for understanding actions made by others (see Introduction). Therefore, the degree to which a region is sensitive to variation in position may provide a hint about the region’s involvement in action performance. To that end, trials were tagged according to the location in which the clip appeared regardless of clip identity, resulting in 49 *t*-values per voxel in each run. The correlation matrix for each ROI consisted of 49×49 values, where each row corresponds to correlations between the pattern (a vector of *t*-values) elicited by various stimuli in one location (in one-half of the data) and the 49 patterns of all locations in the other half (Fig. 3*a,b*). The 49 correlation values in each single row can be presented as a 7×7 matrix termed here the correlation map. This is a spatial representation of the correlation between the response pattern to that location (in one-half of the data) with the patterns of response to all other locations (in the other half of the data; averaged across different splits of the data). Figure 3*c*, bottom, depicts an example of such a correlation map in each of the ROIs. In general, the correlation decreases as the distance between the two locations of the viewed clips on the screen is greater; however, it decreases at a slower rate in LOC and parietal regions than in EVC. This is probably related to the smaller population receptive field (pRF) sizes in EVC compared with the higher-level regions. Prior studies in monkeys have clearly indicated that the receptive fields become larger along the cortical hierarchy in both the ventral and dorsal pathway (Blatt et al., 1990).

Both correlation and classification methods show that EVC is most sensitive to stimulus position, whereas aSPL is the least sensitive and is largely position invariant (Figs. 3*e*, 4*c*). Indeed, results from both methods reveal that aSPL localization performance is significantly lower than both pIPS and LOC ($p < 0.0001$ for both methods), although aSPL still con-

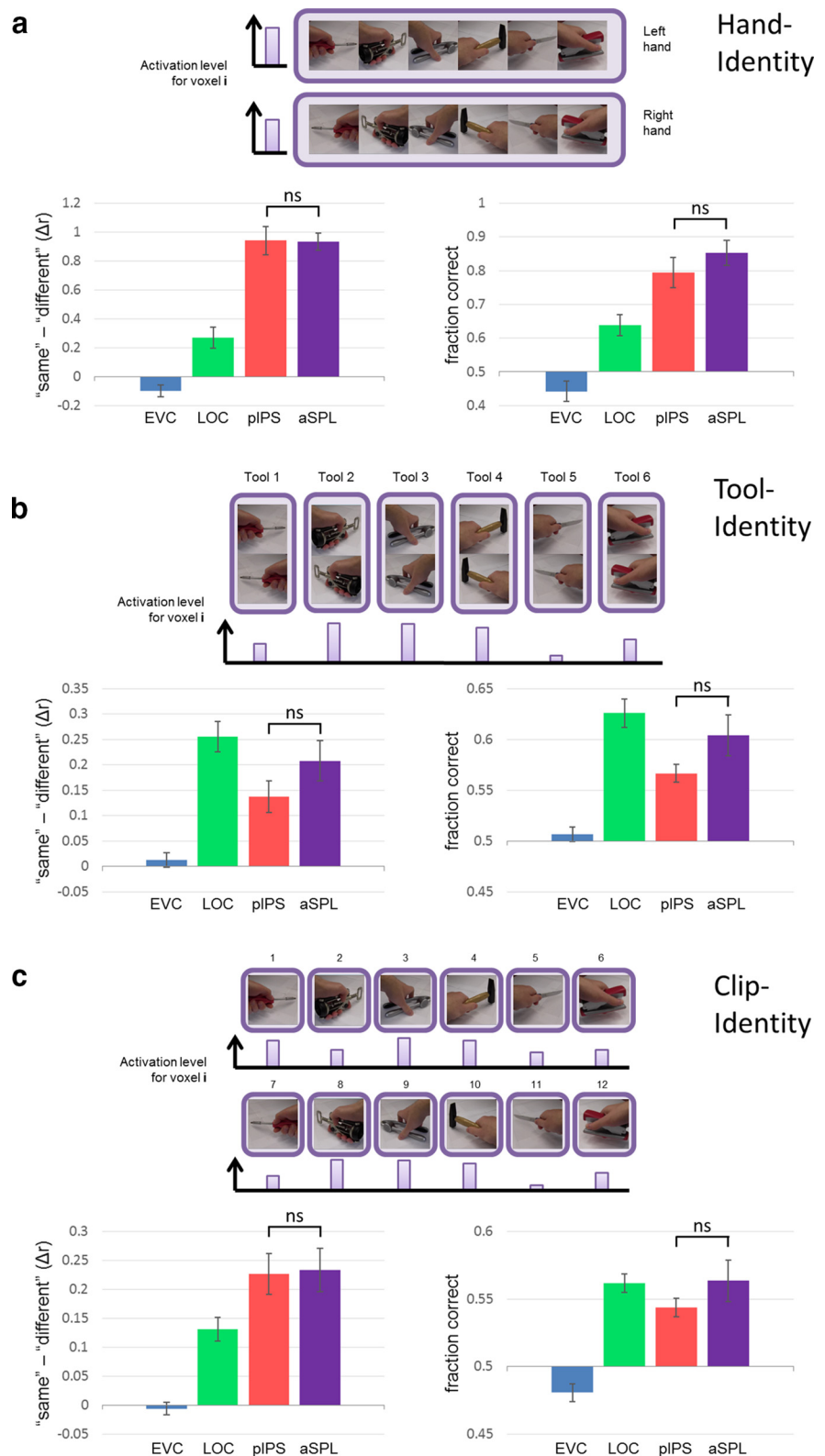


Figure 5. Identity analyses results. Left and right graphs depict results of the correlation and classification methods, respectively. *a*, Hand identity analyses. Top, Trials were divided into two conditions (right hand, left hand), resulting in two coefficients for each voxel. *b*, Tool identity analyses. Top, Trials were divided into six conditions (for the six different tools), resulting in six coefficients for each voxel. *c*, Clip identity analyses. Top, Trials were divided into 12 conditions (for all 12 different combinations of hand and tool), resulting in 12 coefficients for each voxel.

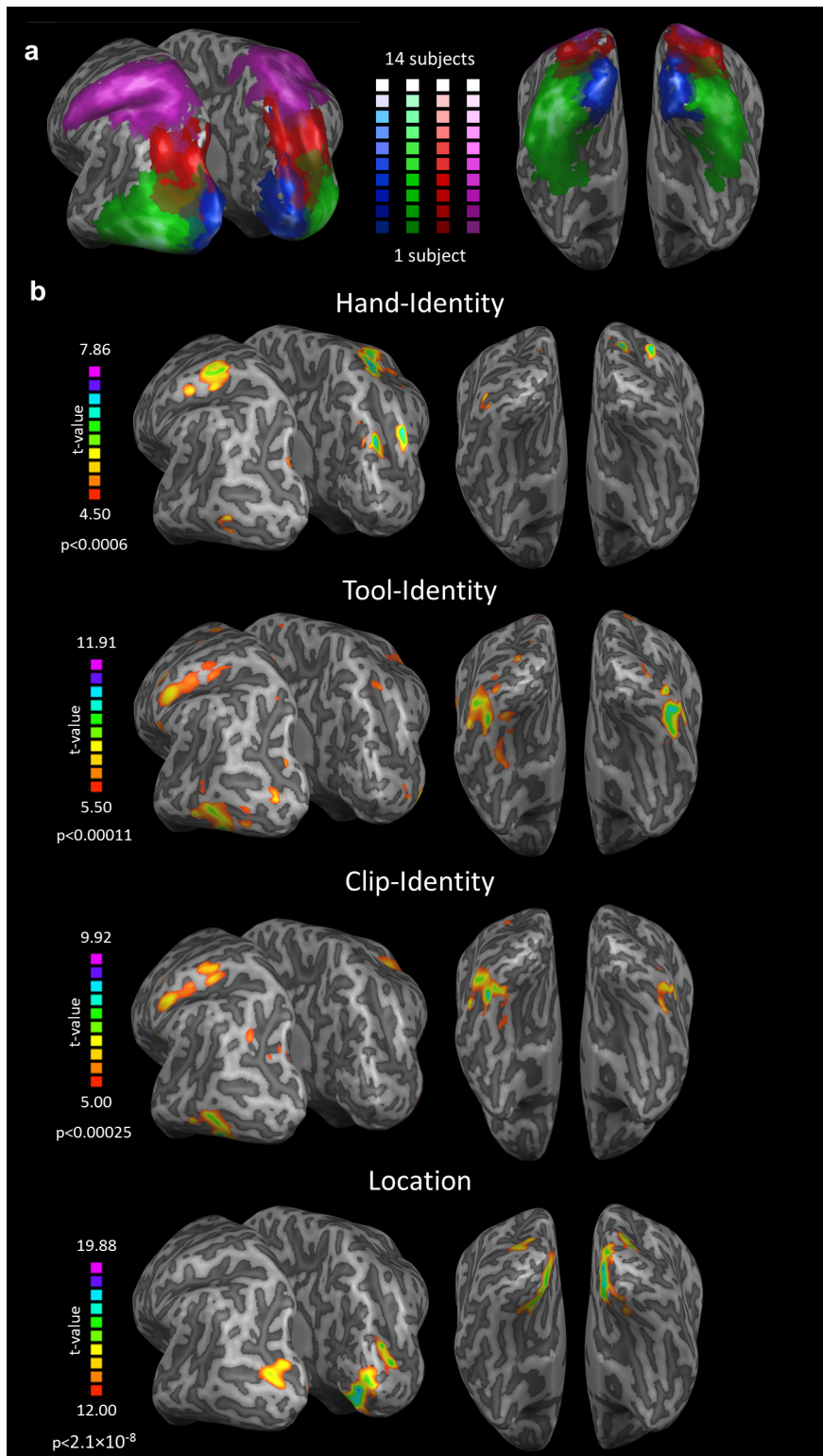


Figure 6. Searchlight results. **a**, Group ROI probability map. All subjects' ROIs were combined and are presented as a single ROI for each region. Brighter colors represent areas (in Talairach coordinates) that are included in a higher number of subjects' ROIs. **b**, Searchlight results (searchlight size: $5 \times 5 \times 5$ voxels). The correlation method was performed for each subject for each voxel throughout the brain. Voxels showing the highest information levels across subjects for a specific characteristic (identity or location) are presented here (color indicates information level) overlaid on one subject's inflated cortex. Left, Dorsolateral view. Right, Ventral view. Thresholds were chosen to present clusters without many small isolated areas and the colors were scaled separately for each of the four analyses.

tains some position information (i.e., above random performance; $p < 0.001$ for both methods). Note that the relatively low amount of positional information using the SVM technique is probably directly related to the random assignment of stimulus location to one of the two classes. Two neighboring locations are often assigned to different classes, requiring discrimination between adjacent locations, which may be beyond the capabilities of the voxel population of aIPS (see Fig. 4a for examples of classifications). In other words, aSPL may indeed carry location information, but at a coarser scale than we tested so far. To test this hypothesis, we tested performance in 20 ordered classifications at five different spatial scales (Fig. 4d). Generally, across ROIs, the pattern of results (Fig. 4e) was similar to the one obtained using a random assignment classification (Fig. 4c). Furthermore, this analysis revealed that aSPL carries location information primarily for coarse-scale classifications (Fig. 4e): decoding accuracy was significantly higher than chance only for the coarsest and second-most-coarse groups of classifications (uncorrected for multiple comparisons; Fig. 4f).

In sum, the patterns of activation to viewed action in LOC are position dependent, although they are much less so than in EVC. This has been noted before in studies using object stimuli (Sayres and Grill-Spector, 2008; Kravitz et al., 2010; Cichy et al., 2013). Here, using many different locations coupled with MVPA, we have been able to give a more detailed account of the different spatial scales at which those regions code stimulus location.

Next, we analyzed the different patterns evoked by the various action clips seen in order to understand what action-specific information is available in the ROIs. Specifically, we wanted to know what degree of information is present regarding hand identity, tool identity, and the combination of both (i.e., clip identity) regardless of the position of the clip on the screen.

Hand identity analysis

Trials were tagged according to which hand (left or right) appeared in each clip regardless of clip location, resulting in two *t*-values (i.e., activation coefficients for the right and left hand clips, respectively) per voxel in each run (Fig. 5a, top). We then tested in each ROI how well patterns of activity can differentiate between clips showing right versus left hand action (grasping a tool) regardless of the location of the clip on the screen

or the specific tool being grasped. Both MVPA methods reached similar results, showing lower performance in LOC than both aSPL and pIPS (correlation method: $t_{(13)} > 6$, $p < 0.0001$; classification method: $t_{(13)} > 3$, $p < 0.05$; Fig. 5a, bottom). Clips of the grasping action performed by the left versus right hand usually differ in the overall position of the hand and therefore may have very different local contrast. However, because in our case each hand clip was seen in all 49 locations, retinal-specific EVC showed no information regarding which hand was active in each clip. The above results, showing clear selectivity to hand identity in the parietal cortex, agree with a previous study finding hand selectivity in aSPL based on the average activation level (Shmuelof and Zohary, 2006).

Tool identity analysis

Trials were tagged according to the tool that appeared in each specific clip. This resulted in six tool-specific activation levels (t -values) per voxel in each run (Fig. 5b, top). In this analysis, we tested how well the activation patterns can differentiate between clips depicting a hand grasping different tools regardless of the clip location and the grasping hand (left or right). LOC showed the highest level of tool specific information, with IPS regions carrying moderate amounts. EVC carries no tool information (across locations). Both MVPA methods reached similar results (Fig. 5b, bottom). The results of this analysis are similar to the previous (hand) analysis in that higher-order regions contain more information about the clip specifics than low-level EVC. Note, however, that although LOC contained significantly lower hand information than both dorsal regions, it showed the highest level of tool information (although this was not significantly different from aSPL).

Clip identity analysis

Trials were tagged according to the specific combination of tool and hand identity that appeared in each clip, resulting in 12 t -values per voxel in each run (Fig. 5c, top). The results based on the correlation method are very similar to those obtained in the hand identity analysis: namely, clip information in LOC is significantly lower than in both parietal regions (pIPS: $t_{(13)} = 2.76$, aSPL: $t_{(13)} = 3.06$, $p < 0.05$ for both). The classification method, however, yielded results that better match the result of the tool analysis (i.e., high accuracy in LOC; Fig. 5c, bottom). We discuss the

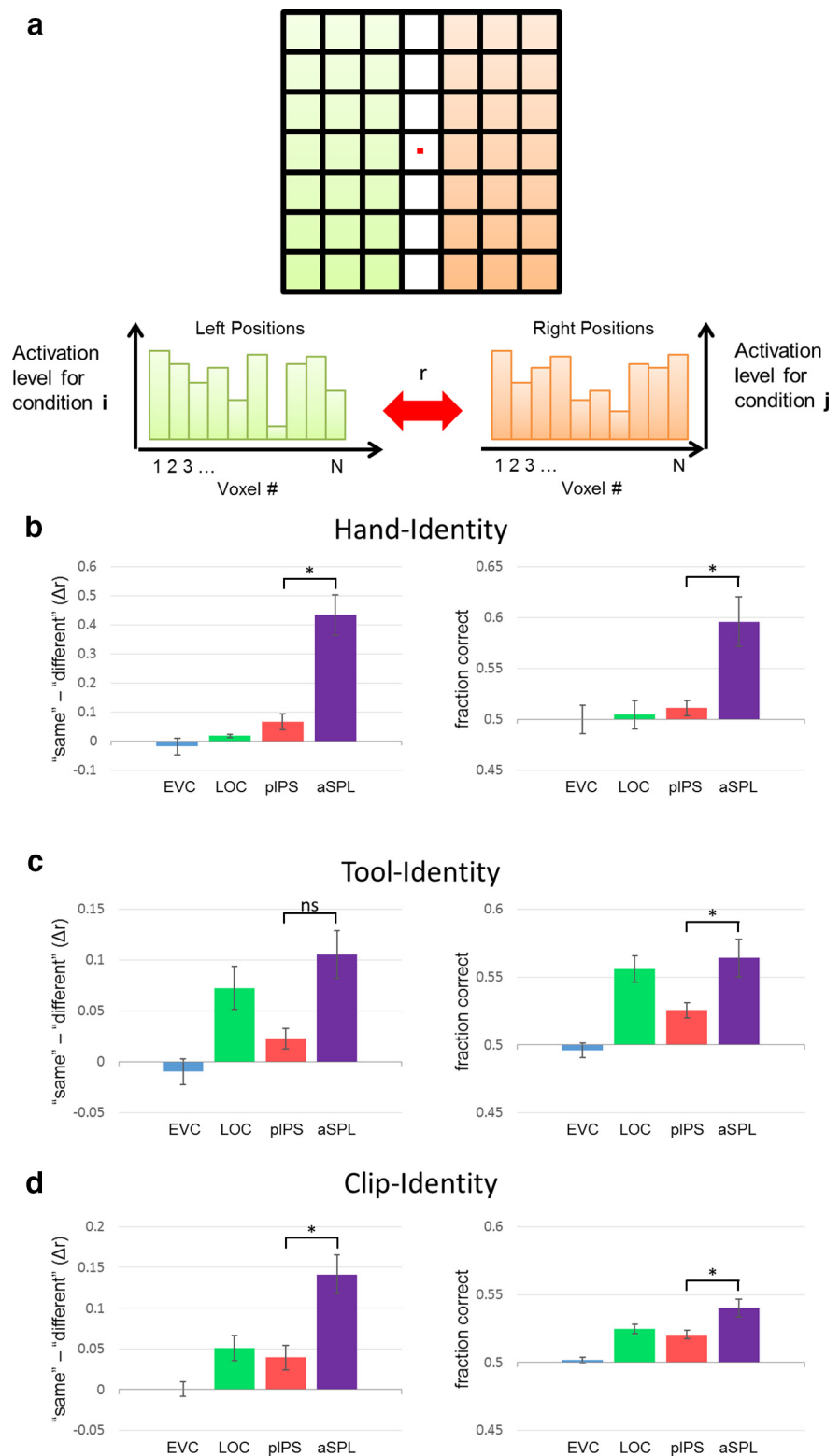


Figure 7. Generalization of identity information across locations. **a**, For the generalization analyses using the correlation method, instead of correlating patterns generated from one-half of the data with patterns from the other half, we correlated patterns from the left hemifield (green squares) with patterns from the right hemifield (orange squares) while ignoring data from the vertical meridian (white squares). For the classification method, we analogously trained the classifier on data from the left hemifield and tested it on the right hemifield (and vice versa). **b–d**, Hand, tool, and clip identity results, respectively. Left, Correlation method. Right, Classification method. Location invariant representations in aSPL generalize well across hemifields.

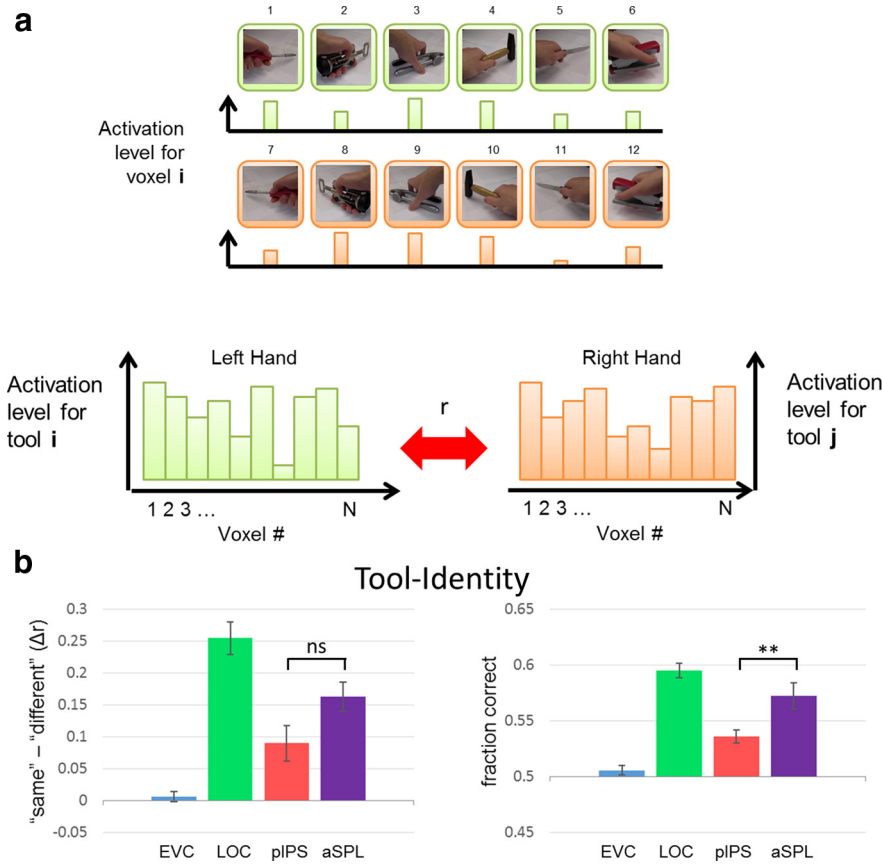


Figure 8. Generalization of tool identity information across hands. **a**, To test for generalization of tool identity across hands, we correlated patterns of response to tools grasped by the left hand (top row; green squares) with patterns of response to tools grasped by the right hand (orange squares). For the classification method, we analogously trained the classifier on data from left-hand clips and tested it on the right-hand clips (and vice versa). **b**, Left, Correlation method results. Right, Classification method results. Both methods yielded results similar to the tool identity analysis (Fig. 5*b*), suggesting that sensitivity to tool identity was independent of hand identity.

possible reasons for this slight divergence later (end of Results section).

Searchlight analysis

The results so far are based on ROI analysis. However, although this analysis enabled us to focus on ROIs that are relevant to location and identity information, we may have missed other relevant regions that were not activated during the independent functional localizer. We therefore performed a whole-brain searchlight analysis using the correlation method to verify our ROI analysis and to determine whether we missed any additional important brain regions. The searchlight analysis results correspond well to the ROI analyses and did not reveal any additional loci carrying information about the hand, tool, or clip identity beyond our choice of ROIs (Fig. 6*b*).

Generalization capabilities

The results so far show that parietal regions contain stimulus identity information across different locations. However, it is of interest whether and to what degree this information is independent of position in the visual field (Rust and Dicarlo, 2010). We hypothesized that areas showing low discriminability for location (such as aSPL) are likely to show the same pattern of activity regardless of the position of the specific image in the visual field. In other words, they would display generalization capabilities. To test this hypothesis, we assessed the patterns of activation for each

image category across all locations in one hemifield and tested the degree to which they matched the patterns of activation for the same stimulus across all locations in the opposite hemifield (Fig. 7*a*). Specifically, we correlated the patterns corresponding to different stimulus identities presented in the LVF with the patterns corresponding to the same stimuli when presented in the RVF and vice versa (Fig. 7*b–d*, left). We also performed a classification analysis in which we trained SVMs to classify the images based on the data from the LVF and tested classification capability of the same images when presented at the RVF, and vice versa (Fig. 7*b–d*, right). As expected, aSPL showed the highest degree of generalization across visual fields for all three identity categories. Therefore, action representations in aSPL are largely invariant to location. Note that although pIPS exhibits identity discriminability across locations (Fig. 5) on par with aSPL, it displays significantly lower generalizability of identity information (Fig. 7). This probably reflects the higher sensitivity to location in pIPS compared with aSPL.

Analogously, we investigated whether the tool information evident in LOC, pIPS, and aSPL was dependent on the identity of the manipulating hand. Specifically, given the high level of hand identity information in the two parietal regions and the dominance of hand information in their correlation matrices, we wondered whether these regions carried any

hand-independent tool identity information. To test this, we compared the patterns of activity elicited by viewing tools manipulated by one hand with patterns of activity for the same tools when manipulated by the other hand (collapsed across all locations; Fig. 8*a*). The pattern of results across ROIs (Fig. 8*b*) was very similar to the original tool identity analysis (in which the data were collapsed across both hands; cf. Fig. 5*b*). This indicates that tool identity information is not driven primarily by the hand similarity between identical clips (i.e., same hand and same tool), but rather by information regarding the tool identity per se (see Discussion).

Comparing correlation and classification analysis

Correlation and classification methods are two different forms of MVPA. Both assess the information available in the patterns of activity of voxel populations, but they focus on somewhat different aspects in the patterns. The classification method handles the data on a single-trial basis, testing the classification on each trial separately. Conversely, the correlation method ignores individual trial variability by pooling across trials in an effort to derive a more reliable assessment of the activation level per condition (e.g., a single coefficient for each condition in each half of the data). Second, and more important, the two methods differ in the weights assigned to each voxel in the pattern. Whereas for computing correlations all voxels are weighted equally, classification with SVM is based on assigning different weights to specific vox-

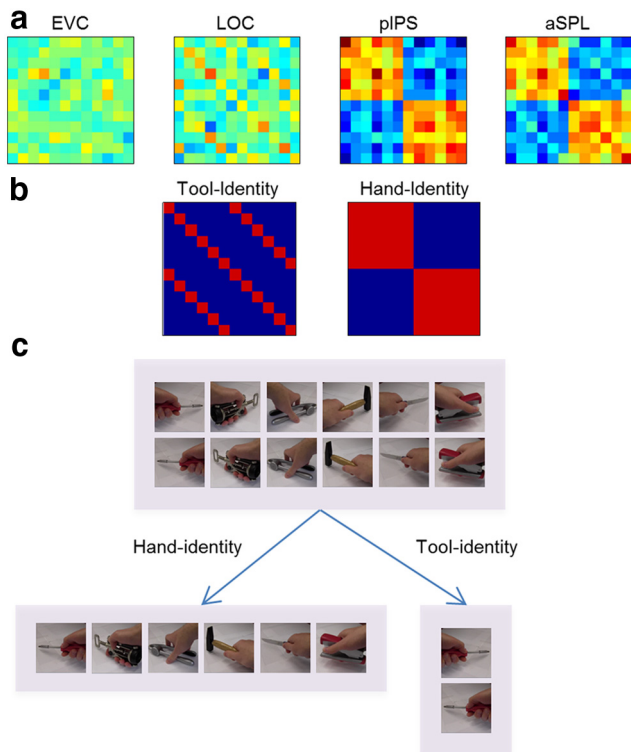


Figure 9. Comparing correlation and classification methods for clip identity. *a*, Clip identity correlation matrices for each ROI averaged across all subjects' individual matrices. Clips are numbered as in Figure 5*c*. Notice that, in pIPS and aIPS, the matrix structure is dominated by hand identity (clips 1–6 vs clips 7–12), whereas in LOC, the structure is dominated by tool identity. *b*, Schematic matrices illustrating extreme cases of complete tool identity information with no hand identity information (left) or complete hand identity information with no tool identity information (right). Notice in *a* that the correlation matrix in LOC is more similar to the left matrix in *b* and, in IPS regions, the matrix is more similar to the right matrix. *c*, Illustration of the effect of perfect hand or tool identity information on classification accuracy. In contrast to the correlation method, tool identity has a stronger effect on classification accuracy than does hand identity. Whereas perfect hand identity information leaves us with six possibilities (bottom left), perfect tool identity information limits us to only two possibilities (bottom right).

els according to their usefulness in classification. Therefore, even if only a few voxels within a specific ROI (typically containing hundreds of voxels) show some differential selectivity to images from the two classes to be differentiated, a successful SVM classifier will assign large weights to those voxels and manage to classify well, whereas the correlation method will not be successful because of the dominance of uninformative voxels that contribute noise, thereby masking the similarity (e.g., correlation) between images of the same category. Despite these differences, the two methods generally yielded similar results. The relative performance per analysis showed a similar pattern (i.e., rank order) across regions for the two methods (Figs. 5*a,b*, 7*b,c*). However, in our clip analysis (which takes into account both hand and tool identity; Fig. 5*c*), the results of the two methods differed: the rank order of the correlation results (between ROIs) was similar to those obtained in the hand analysis (larger amount of information in parietal regions, smaller in LOC; Fig. 5*a*), whereas the classification results were similar to the ones in the tool analysis (greater information in LOC and aSPL compared with pIPS; Fig. 5*b*). This may be a consequence of the larger number of tool analysis conditions (six) relative to the hand analysis conditions (two), the specific combinations of which determine the clip identity conditions (12), combined with the higher correlation information in the hand analysis relative to the tool analysis (cf.

values in Fig. 5*b,c*). As a result, hand identity dominates the correlation values in the clip identity analysis (this is vividly seen in the correlation matrices in Fig. 9*a*). Indeed, the correlation results across ROIs in the clip identity analysis match closely the hand identity results. Conversely, because reliable classification requires accurate separation between the representations of all 12 clips, the larger number of tool conditions means that tool identity information should dominate classification accuracy over hand identity information. To illustrate this, if a multiclass classifier (i.e., classifies each clip as one of 12 options) is able to determine perfectly which tool is grasped in the clip, it will have 50% accuracy (where the chance level is 1 of 12, or 8.3%), having to guess only the hand identity. In contrast, if the classifier knows only the hand identity, it will have to guess the tool (out of six options), resulting in 16.6% accuracy (Fig. 9*c*). In other words, given our specific configuration of stimuli, tool identity should have a higher impact on clip identity classification than hand identity. It therefore makes perfect sense that, when using the correlation method, the clip identity results resemble the hand identity results, whereas with the classification method, the clip analysis results are more similar to the tool analysis.

Relationship between information and tool action representation in visual cortex

We have shown that there is potential information regarding the position and identity of tool manipulation clips in the patterns of activation in various regions of the visual cortex. For example, the patterns in both LOC and aSPL allow reliable decoding of tool identity. Does this imply that there are similarities between the representations of viewed actions in the ventral and dorsal visual streams in terms of their mean level and/or distribution of BOLD activation to the different stimulus conditions? To address this, we calculated, for each ROI, the mean *t*-values for all the relevant conditions and also a few examples of the multivoxel activation patterns for various stimulus locations. The results are shown in Figure 10. Within each ROI, the differences between activation levels across conditions were minute and do not provide enough information to differentiate between various conditions. In LOC, pIPS, and aIPS, the mean *t*-values were positive for both identity and location analysis (Fig. 10*a*, top and bottom, respectively). In EVC, the mean *t*-values were mostly negative, except for the most central positions, which were positive. This can also be seen at the single-voxel level (Fig. 10*b*). The negative values in EVC for the identity analysis (hand, tool, or clip) were a result of presenting stimuli in peripheral locations in the majority of cases. Voxels in EVC have center-surround pRFs and the size of the excitatory center pRF is small relative to the eccentricities used in our study (Zuiderbaan et al., 2012). For most voxels, the stimulated location is usually outside the positive region of the pRF and they therefore either have no BOLD response (i.e., a *t*-value ~ 0) or have a negative response when the stimulus is located in their negative surround region. In any case, previous MVPA studies have assumed that potential information corresponds to “representation” regardless of whether the BOLD responses of individual voxels were positive or negative. The standard has been to use a normalized measure of the voxel's activation [e.g., the difference in activation level from the mean level of activation across all stimuli (Haxby et al., 2001) or *t*-value, scaling this difference by the voxel's variation in the null period (Kriegeskorte et al., 2008) as the voxel's contribution to the pattern vector]. Clearly, to prove the existence of a neural representation, it would be necessary to show a causal link to perception; that is, to create a certain pattern of

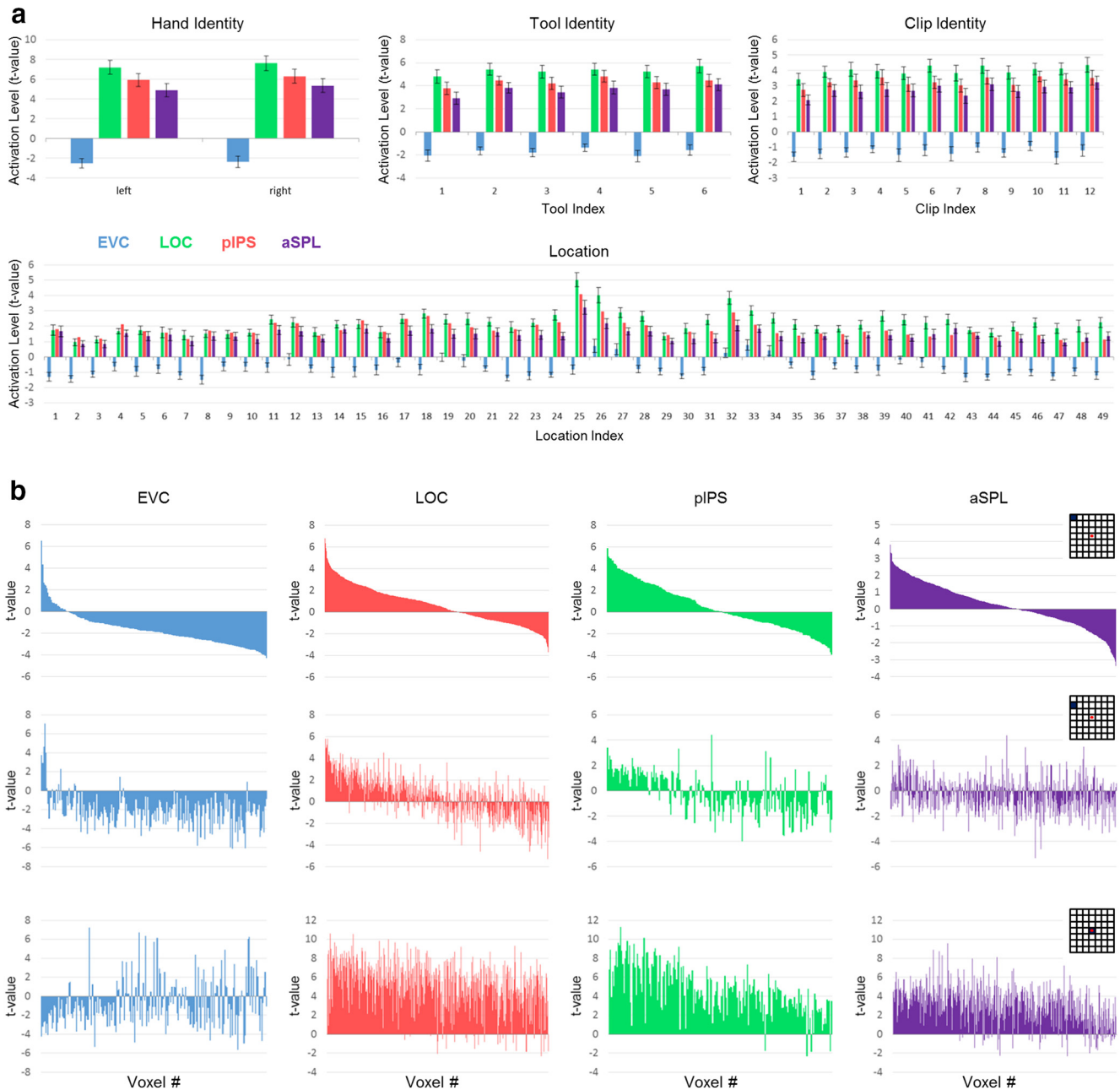


Figure 10. Mean ROI activations for the various stimulus conditions and examples of the multivoxel activation patterns. **a**, Mean *t*-values per ROI and conditions for each identity and location analysis. Bar color denotes the ROI (blue: EVC; green: LOC; red: pIPS; purple: aSPL). For the tool identity and clip identity analyses, conditions are numbered in accordance with Figure 5, *b* and *c*. For the location analysis, conditions are numbered according to the order depicted in Figure 3*b*. Note that, within each ROI, the differences between activation levels across conditions (depicted by the same color) are minimal and do not provide enough information to differentiate between the various conditions. **b**, Examples of patterns of activation for specific stimulus locations. Vectors of *t*-values across all voxels are shown for each ROI in response to three different locations for a single subject. The locations shown are (from top to bottom) #1, #2, and #25, which are represented schematically in the rightmost column. Voxels are sorted according to their *t*-values in response to the first location. Note that in EVC, LOC, and pIPS, the vectors in response to distant locations are quite different from one another (e.g., cf. first row with third row).

activity (e.g., by mean of optogenetics or focused stimulation) and show that it causes a specific perception.

Discussion

Summary of results

We characterized the information available in the patterns of fMRI activation when viewing object-grasping actions in various cortical regions. Our results suggest that in a functional gradient within the dorsal visual stream along the posterior–anterior axis, position information is gradually lost, whereas hand and tool

identity information is maintained. This may reflect a transformation of visual input from a highly specific retinotopic representation in early visual areas to an abstract, position invariant representation of viewed actions in the aSPL.

Limitations of our experimental design

We found that the patterns of activation carry tool identity information in several cortical regions. However, the variation in the patterns elicited by the various tools may be due to different causes. One possibility is that, because different tools are grabbed

in diverse ways, they may evoke various tool affordances; however, the hand generalization analysis suggests that this is not a primary factor because clips with opposite hand identity are somewhat different in their affordances (in terms of the angle at which the tool is grasped). Another possibility is that the typical use (e.g., manipulation) varies considerably between the tools used in the current experiment. We cannot tease apart these two possible sources of discriminating information in the current study. It is obviously also possible that the information in some regions may reflect tool identity per se. Similarly, areas that contain information about hand identity may actually encode hand motion direction because the right-hand clips always depict motion in the leftward direction and vice versa.

Another issue is the degree to which the information is relevant to the task at hand. Task requirements may shape the sensitivity to the specific parameters of the stimuli (e.g., location and identity). In the current task, subjects were asked to name covertly the identity of the clips regardless of their location. Instructing subjects to direct their attention to the position of the stimulus on the screen while ignoring its identity may change the activation patterns, possibly resulting in greater location information and a lower degree of identity information across the various ROIs.

Finally, because activity in the parietal cortex is known to be modulated by attention, it is likely that the location information that we found in the parietal cortex reflects not only the stimulus location, but also the location to which the subject attended.

Relevance to previous MVPA studies

In recent years, several studies have used MVPA to assess the information available regarding the identity of viewed actions in various areas of the dorsal stream. Dinstein et al. (2008) presented subjects with clips of hand actions and found that the action identity (“rock,” “paper,” or “scissors”) can be reliably assessed (i.e., above chance level) from the pattern of BOLD activity in anterior IPS, which partially overlaps our aSPL ROI. However, a much better decoding accuracy was found in EVC, suggesting that differences in low-level visual features (e.g., variations in the retinal image) are likely to be a major factor determining the decoding level. More recently, Ogawa and Inui (2011) had subjects view static images of hand–object interactions. Applying SVMs to the patterns of responses evoked by each image in various ROIs, they attempted to classify each image according to specific properties. Decoding accuracy in aIPS was above chance for action type, object identity, and hand “side.” However, because all stimuli were shown at fixation, most classifications were confounded by low-level differences and, indeed, decoding accuracy remained high when using vectors of stimuli pixelwise luminance values. Nevertheless, their final analysis, which involved decoding the action type while generalizing across two other properties, yielded significant results in aIPS and premotor cortex, but not in EVC. This result presumably reflects coding of high-level visual action features rather than low-level visual features.

Our results are consistent with these previous studies and extend them by showing that the information in anterior parietal cortex is not limited to low-level retinotopic differences (as verified by low decoding accuracy of stimulus identity in EVC), but rather contains high-level information regarding action properties (i.e., identity of the hand and tool involved in grasping).

Similarly, Wurm and Lingnau (2015) have shown recently that activation patterns in the inferior parietal lobule (IPL, adjacent to aSPL) and lateral occipitotemporal cortex (LOTc, a re-

gion in LOC) enable decoding the viewed action (i.e., opening or closing) across different action kinematics (i.e., twisting or pulling/pushing) and object categories (i.e., bottle or jar). Conversely, EVC does not show significant decoding, indicating that the generalization properties in IPL and LOTc do not rely on low-level features.

Similarities between ventral and dorsal stream representations

One interesting aspect of our study is that there are some similarities between the representations of viewed actions in the ventral and dorsal visual streams. This has also been noted in past studies. For example, using fMRI repetition suppression, Konen and Kastner (2008) showed that LOC and pIPS regions (IPS1 and IPS2) both represent objects in a view- and size-invariant manner. Using surface-based searchlight MVPA, Oosterhof et al. (2010) found that, in both LOC and a region approximately corresponding to aSPL, observed and executed actions have similar vector representations. Bracci and Peelen (2013) found that, in both IPS and left LOTc, representations of tool images reflect the degree to which they serve as extensions of the body or hand. These fMRI findings, as well as recent functional (Mahon et al., 2013) and anatomical (Takemura et al., 2015) connectivity studies, indicate that, although the two streams may have different functional preferences (Goodale and Milner, 1992; Shmuelof and Zohary, 2005), significant communication channels exist between ventral and dorsal visual regions.

Ventral gradient of location and identity information

A large number of studies has uncovered a posterior–anterior gradient along the ventral stream in which low-level visual input is transformed into high-level object representations (Grill-Spector and Malach, 2004). For example, it has been shown that LOC is less sensitive than EVC to the contrast level in an image (Avidan et al., 2002) and also shows a large degree of invariance to changes in position or retinal size while at the same time being viewpoint selective (Grill-Spector et al., 1999). Our results are consistent with this ventral gradient, showing that representations in LOC are less dependent on stimulus location than the posterior EVC and reflect aspects such as tool identity.

Dorsal gradient of location and identity information

It has been suggested that a visual–somatic (Blangero et al., 2009) or a visual–motor (Stark and Zohary, 2008) gradient exists along the anterior–posterior axis of the dorsal stream (see Heed et al., 2011 for a similar eye–hand motor gradient). Specifically, during action execution (such as grasping or reaching), posterior regions in the IPS are more involved in processing relevant visual aspects (such as target position), whereas anterior regions are more involved in encoding of the motor aspects (i.e., identity of the hand performing the action). The same gross division can also be seen here for purely visual tasks. Our results show that location information is gradually lost along the posterior–anterior axis, thus extending our previous results (Porat et al., 2011), whereas representations become increasingly sensitive to tool and hand identity even when no action is performed by the observer.

Abstract representations in the ventral and dorsal streams

We have shown that there are commonalities in the representation of viewed actions in LOC and parietal cortex: both show clear feature generalization compared with EVC. This raises the question regarding the division of labor between the dorsal and ventral streams. We found that the representation in LOC was

still largely location dependent, whereas aSPL showed a location-invariant activation pattern. Furthermore, LOC carried less hand identity information (across all locations) than aSPL. These results may be interpreted as reflecting a more abstract representation in the dorsal stream than in the ventral. However, it is possible that the representations that we have identified do not reflect the final word of either stream and that more anterior regions may host more abstract representations. Such hypothetical abstract representation may generalize across tools and hands, showing instead selectivity to the action, such as “cutting,” “throwing,” or “grasping.” Because all of the clips that we used depict grasping, such an area would not differentiate between the different clips, instead generalizing across the different tools and hands. Consistent with the ventral stream abstraction gradient discussed above, the anterior temporal lobe (ATL) may be a reasonable candidate for hosting a highly abstract representation because it has been suggested that a gradient of conceptual information in the ventral stream culminates in ATL, where activity patterns reflect the actions associated with certain tools (e.g., kitchen vs garage tools; Peelen and Caramazza, 2012).

One way to investigate the “goal” of each stream is by studying the stimulus representations and transformations taking place along each stream. Here, we analyzed the representations with regard to stimulus position and stimulus identity. Future studies may incorporate a similar approach with regard to additional features such as viewpoint and gaze direction. It remains to be determined whether the degree of invariance and selectivity to other visual properties align with the posterior–anterior gradient that we have suggested.

References

- Avidan G, Harel M, Hendler T, Ben-Bashat D, Zohary E, Malach R (2002) Contrast sensitivity in human visual areas and its relationship to object recognition. *J Neurophysiol* 87:3102–3116. [Medline](#)
- Blangero A, Menz MM, McNamara A, Binkofski F (2009) Parietal modules for reaching. *Neuropsychologia* 47:1500–1507. [CrossRef Medline](#)
- Blatt GJ, Andersen RA, Stoner GR (1990) Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *J Comp Neurol* 299:421–445. [CrossRef Medline](#)
- Bracci S, Peelen MV (2013) Body and object effectors: the organization of object representations in high-level visual cortex reflects body-object interactions. *J Neurosci* 33:18247–18258. [CrossRef Medline](#)
- Chang C-C, Lin C-J (2011) LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)* 2:27.
- Cichy RM, Sterzer P, Heinzle J, Elliott LT, Ramirez F, Haynes JD (2013) Probing principles of large-scale object representation: Category preference and location encoding. *Hum Brain Mapp* 34:1636–1651. [CrossRef Medline](#)
- Dinstein I, Gardner JL, Jazayeri M, Heeger DJ (2008) Executed and observed movements have different distributed representations in human aIPS. *J Neurosci* 28:11231–11239. [CrossRef Medline](#)
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. *Neuroimage* 7:30–40. [CrossRef Medline](#)
- Garrido L, Vaziri-Pashkam M, Nakayama K, Wilmer J (2013) The consequences of subtracting the mean pattern in fMRI multivariate correlation analyses. *Front Neurosci* 7.
- Golomb JD, Kanwisher N (2012) Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cereb Cortex* 22:2794–2810. [CrossRef Medline](#)
- Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:20–25. [CrossRef Medline](#)
- Grill-Spector K, Malach R (2004) The human visual cortex. *Annu Rev Neurosci* 27:649–677. [CrossRef Medline](#)
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R (1999) Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24:187–203. [CrossRef Medline](#)
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430. [CrossRef Medline](#)
- Heed T, Beurze SM, Toni I, Röder B, Medendorp WP (2011) Functional rather than effector-specific organization of human posterior parietal cortex. *J Neurosci* 31:3066–3076. [CrossRef Medline](#)
- Konen CS, Kastner S (2008) Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci* 11:224–231. [CrossRef Medline](#)
- Kravitz DJ, Kriegeskorte N, Baker CI (2010) High-level visual object representations are constrained by position. *Cereb Cortex* 20:2916–2925. [CrossRef Medline](#)
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60:1126–1141. [CrossRef Medline](#)
- Mahon BZ, Kumar N, Almeida J (2013) Spatial frequency tuning reveals interactions between the dorsal and ventral visual systems. *J Cogn Neurosci* 25:862–871. [CrossRef Medline](#)
- Ogawa K, Inui T (2011) Neural representation of observed actions in the parietal and premotor cortex. *Neuroimage* 56:728–735. [CrossRef Medline](#)
- Oosterhof NN, Wiggett AJ, Diedrichsen J, Tipper SP, Downing PE (2010) Surface-based information mapping reveals crossmodal vision–action representations in human parietal and occipitotemporal cortex. *J Neurophysiol* 104:1077–1089. [CrossRef Medline](#)
- Peelen MV, Caramazza A (2012) Conceptual object representations in human anterior temporal cortex. *J Neurosci* 32:15728–15736. [CrossRef Medline](#)
- Porat Y, Pertzov Y, Zohary E (2011) Viewed actions are mapped in retinotopic coordinates in the human visual pathways. *J Vis* 11:17. [Medline](#)
- Roth ZN, Zohary E (2014) Fingerprints of learned object recognition seen in the fMRI activation patterns of lateral occipital complex. *Cereb Cortex*. In press.
- Rust NC, Dicarlo JJ (2010) Selectivity and tolerance (“invariance”) both increase as visual information propagates from cortical area V4 to IT. *J Neurosci* 30:12978–12995. [CrossRef Medline](#)
- Sayres R, Grill-Spector K (2008) Relating retinotopic and object-selective responses in human lateral occipital cortex. *J Neurophysiol* 100:249–267. [CrossRef Medline](#)
- Shmuelof L, Zohary E (2005) Dissociation between ventral and dorsal fMRI activation during object and action recognition. *Neuron* 47:457–470. [CrossRef Medline](#)
- Shmuelof L, Zohary E (2006) A mirror representation of others’ actions in the human anterior parietal cortex. *J Neurosci* 26:9736–9742. [CrossRef Medline](#)
- Stark A, Zohary E (2008) Parietal mapping of visuomotor transformations during human tool grasping. *Cereb Cortex* 18:2358–2368. [CrossRef Medline](#)
- Takemura H, Rokem A, Winawer J, Yeatman JD, Wandell BA, Pestilli F (2015) A major human white matter pathway between dorsal and ventral visual cortex. *Cereb Cortex*. In press.
- Valyear KF, Culham JC, Sharif N, Westwood D, Goodale MA (2006) A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: a human fMRI study. *Neuropsychologia* 44:218–228. [CrossRef Medline](#)
- Weber M, Thompson-Schill SL, Osherson D, Haxby J, Parsons L (2009) Predicting judged similarity of natural categories from their neural representations. *Neuropsychologia* 47:859–868. [CrossRef Medline](#)
- Wurm MF, Lingnau A (2015) Decoding actions at different levels of abstraction. *J Neurosci* 35:7727–7735. [CrossRef Medline](#)
- Zuiderbaan W, Harvey BM, Dumoulin SO (2012) Modeling center–surround configurations in population receptive fields using fMRI. *J Vis* 12:10. [Medline](#)