

Disentangling Representations of Object and Grasp Properties in the Human Brain

 Sara Fabbri,^{1,2,4}  Kevin M. Stubbs,¹  Rhodri Cusack,^{1,2,3} and  Jody C. Culham^{1,2,5}

¹The Brain and Mind Institute, ²Department of Psychology, and ³Department of Medical Biophysics, University of Western Ontario, London, Ontario N6A 5C2, Canada, ⁴Radboud University, Donders Institute for Brain, Cognition, and Behaviour, Nijmegen, 6500 HE Nijmegen, The Netherlands, and ⁵Center for Mind/Brain Sciences, University of Trento, 38068 Mattarello, Italy

The properties of objects, such as shape, influence the way we grasp them. To quantify the role of different brain regions during grasping, it is necessary to disentangle the processing of visual dimensions related to object properties from the motor aspects related to the specific hand configuration. We orthogonally varied object properties (shape, size, and elongation) and task (passive viewing, precision grip with two or five digits, or coarse grip with five digits) and used representational similarity analysis of functional magnetic resonance imaging data to infer the representation of object properties and hand configuration in the human brain. We found that object elongation is the most strongly represented object feature during grasping and is coded preferentially in the primary visual cortex as well as the anterior and posterior superior-parieto-occipital cortex. By contrast, primary somatosensory, motor, and ventral premotor cortices coded preferentially the number of digits while ventral-stream and dorsal-stream regions coded a mix of visual and motor dimensions. The representation of object features varied with task modality, as object elongation was less relevant during passive viewing than grasping. To summarize, this study shows that elongation is a particularly relevant property of the object to grasp, which along with the number of digits used, is represented within both ventral-stream and parietal regions, suggesting that communication between the two streams about these specific visual and motor dimensions might be relevant to the execution of efficient grasping actions.

Key words: fMRI; grasping; motor; MVPA; representational similarity analysis; visual

Significance Statement

To grasp something, the visual properties of an object guide preshaping of the hand into the appropriate configuration. Different grips can be used, and different objects require different hand configurations. However, in natural actions, grip and object type are often confounded, and the few experiments that have attempted to separate them have produced conflicting results. As such, it is unclear how visual and motor properties are represented across brain regions during grasping. Here we orthogonally manipulated object properties and grip, and revealed the visual dimension (object elongation) and the motor dimension (number of digits) that are more strongly coded in ventral and dorsal streams. These results suggest that both streams play a role in the visuomotor coding essential for grasping.

Introduction

To efficiently grasp objects, the visuomotor system needs to ascertain object properties and use them to prepare the appropriate grasp. For example, a small round sphere, such as a pearl, is

typically grasped between the thumb and the index finger (a grip we denote as Precision 2), while a long object, such as a dumbbell, is typically grasped with the palm and all the five fingers (Coarse 5). To characterize the contribution of different brain regions during grasping, it is necessary to distinguish whether these regions represent object properties, grip type, or both.

Neurophysiological studies have identified two core regions involved in grasping: the anterior intraparietal area (AIP) and premotor area F5 (Taira et al., 1990; Jeannerod et al., 1995; Murata et al., 1997, 2000). However object and grip type covaried in these studies, as the monkey grasped each object with its natural grip. Therefore, it is unclear to what extent neurons in these regions are representing object type, grip type, or both. A few studies have orthogonally manipulated these dimensions and reported that both regions in the monkey contained neurons selec-

Received Jan. 28, 2016; revised June 2, 2016; accepted June 6, 2016.

Author contributions: S.F. and J.C.C. designed research; S.F. and K.M.S. performed research; S.F., K.M.S., R.C., and J.C.C. analyzed data; S.F., R.C., and J.C.C. wrote the paper.

This work was supported by grants to J.C.C. from the Natural Sciences and Engineering Research Council of Canada (Discovery Grant 249877-2006-RGPIN and Collaborative Research and Training Environment Grant 371161-2009 CREAT) and the Canadian Institutes of Health Research (MOP 130345). We thank Dr. Derek Quinlan and Adam McLean for technical assistance.

The authors declare no competing financial interests.

Correspondence should be addressed to Sara Fabbri, Radboud University, Donders Institute for Brain, Cognition, and Behaviour, Montessorilaan 3, 6525 HR, Nijmegen, The Netherlands. E-mail: s.fabbri@donders.ru.nl.

DOI:10.1523/JNEUROSCI.0313-16.2016

Copyright © 2016 the authors 0270-6474/16/367648-15\$15.00/0

tive for grip type, object orientation, and their interaction (Baumann et al., 2009; Fluet et al., 2010), but during movement execution, the majority of cells in the AIP coded both grip type and object orientation (Baumann et al., 2009), while in area F5 the majority coded grip information (Fluet et al., 2010).

The corresponding areas in humans, the anterior intraparietal sulcus (aIPS) and the ventral premotor cortex (PMv), have been found to activate more strongly during grasping than reaching (Culham et al., 2003; Frey et al., 2005; Cavina-Pratesi et al., 2010) and to show selectivity for grip type (Gallivan et al., 2011; Fabbri et al., 2014). However, these neuroimaging studies focused on the motor aspects of the grasping task without manipulating object properties. To our knowledge, the only study that orthogonally manipulated grip type (Precision 2 and Coarse 5) and object size (small and large sphere; Begliomini et al., 2007) found higher activity in the aIPS during Precision 2 than Coarse 5 grasping, regardless of object size, suggesting that the aIPS represents mainly the hand configuration and not object information during grasping. These results have been recently extended with multivariate pattern analysis by Di Bono et al. (2015). This selectivity in the aIPS for grip type and not object properties appears in conflict with sensitivity for object orientation reported in monkey areas AIP and F5 (Baumann et al., 2009; Fluet et al., 2010). An explanation for this conflict may be that orientation selectivity reported in monkeys reflects selectivity for the orientation of the wrist instead of the object. Therefore, it remains unclear whether object properties are coded in grasping regions and which object properties are more relevant than others during grasping.

To disentangle the representation of visual and motor codes for grasping, human participants passively viewed or grasped in various ways (Precision 2, Precision 5, and Coarse 5) six real objects (plates, disks, spheres, cubes, cylinders, and bars) presented in three sizes (small, medium, and large) during functional magnetic resonance imaging (fMRI). We used representational similarity analysis (RSA; Kriegeskorte et al., 2008; Nili et al., 2014) to distinguish the representation of visual (object shape, size, and elongation) and motor (number of digits, degree of precision) aspects of the grasping tasks in regions of the grasping circuits. We also included regions of the ventral stream, as previous studies have shown that ventral-stream area lateral occipital complex (LOC) codes grasp-relevant dimensions during action planning (Gallivan et al., 2014; Monaco et al., 2014). Specifically, we predicted that visuomotor areas, including the aIPS, represent both visual and motor aspects of the grasping task, while motor areas, including the PMv, primarily represent motor codes. Furthermore, we tested which object dimensions (shape, size, or elongation) are most strongly represented during grasping. Finally, given evidence of ventral-stream involvement in grasping, we expected representation of object properties in both dorsal and ventral streams.

Materials and Methods

Participants. Analyses were based on data from 12 neurologically healthy adult participants (eight females and four males) recruited for this study; mean age was 24 years. Due to excessive head motion, data from an additional three participants were excluded. All participants were right-handed and had normal or corrected-to-normal visual acuity. Informed consent was obtained from each participant in accordance with the University of Western Ontario Health Sciences Research Ethics Board and the Declaration of Helsinki.

Objects. In a rapid event-related fMRI design, participants directly viewed a series of 18 objects and performed one of four tasks. The 18 objects were a factorial combination of six shapes and three sizes created from blocks of maple wood and painted white. The six shapes were a factorial combination of three types of elongation (flat, isotropic, or

elongated) and two types of edges (shape: square or round), yielding a square plate, cube, elongated square bar, disk, sphere, or cylinder, as shown in Figure 1*a*. Each object could have one of three sizes along the depth dimension (7.6, 5.7, or 3.8 cm); because subjects were instructed to grasp the objects along this dimension, this corresponded to the final grip aperture size, which has been shown to be particularly relevant in fMRI studies of grasping (Monaco et al., 2014). To vary the elongation, the other dimensions were either equal to the size (depth) dimension or fixed (14 cm width for elongated objects; 0.8 cm height for the planar objects; Fig. 1*a*). Note that elongation here is not strictly equivalent to the aspect ratio of the objects as they would be projected on the retina. Our choice of dimensions was based on two considerations: (1) the aim to vary the graspable dimension (and round/square shape) independently of other dimensions; and (2) the aim to investigate objects that included planar shapes and elongated objects, two categories for which neural subpopulations have been identified using neurophysiology (Sakata et al., 1998), as well as an intermediate (isotropic) shape.

Tasks. The four tasks included passive viewing and three grasping tasks that differed in the number of digits used and the precision required (Fig. 1*a*). In the first grasping task, participants used the index finger and thumb to precisely grasp the object (Precision 2 condition). In the second grasping task, participants used all five digits to precisely grasp the object (Precision 5 condition). In the third grasping task, participants grasped the object coarsely in a whole-hand grasp using all five digits (Coarse 5 condition). All grasps were performed with the right hand.

Apparatus. Participants lay supine in the scanner with the head tilted 20° to allow direct sight of the objects placed in front of them. Each stimulus was placed on a pedestal positioned on a wooden platform above the participant's pelvis (Fig. 1*b*), which was placed such that the objects on the pedestal could be comfortably grasped. Each participant was instructed to place his or her right hand on the platform in a comfortable resting position (Fig. 1*b*) and to maintain fixation during the entire experiment on the light-emitting diode (LED) placed above the object by ~10–15° of visual angle. Participants wore headphones to receive auditory instructions. To limit head movements during grasping, the participant's right forearm was strapped to the bed. Throughout the experiment, the participant's actions were recorded by an MR-compatible infrared camera (MRC Systems) placed at the entrance of the magnet bore. Behind the participant's head, a series of LEDs illuminated the workspace during the trial and turned off during the intertrial interval (ITI).

To efficiently alternate placement of the 18 objects, two experimenters were required. One experimenter stood in front of a table where the 18 objects were each positioned beside an LED that illuminated to indicate when a given object should be passed to a second experimenter. The second experimenter placed the appropriate object for the upcoming trial on the pedestal in darkness during the ITI.

Procedure. Consistent with other RSA studies using many stimuli (Kriegeskorte et al., 2008), we used a fast event-related design to present 10–12 experimental runs (depending on time constraints) in which each combination of 18 objects \times 4 tasks (72 trials) was presented once per run in a pseudorandom order for each run (with the constraint that the same object was not presented more than twice in sequence, even if different actions were performed upon it). At the start of each trial, the workspace was illuminated and participants received a one-syllable auditory instruction of the task to execute: "None" (passive viewing), "Two" (Precision 2), "Five" (Precision 5), or "Palm" (Coarse 5). Participants had 2 s of illumination in which to execute the task with visual feedback and return the hand to the starting position. The windows in the scanner room were covered with black quilts and the room lights were off such that the fixation LED was the only thing visible when the illuminator lights were off during the ITI. Each run lasted 9 min and began and ended with 16 s of a resting baseline in darkness. To decorrelate the hemodynamic responses of successive trials, we jittered the ITI: 2 s in 50% of the trials, 6 s in 33% of the trials, and 8 s in 17% of the trials. The distribution of ITIs was counterbalanced across conditions between runs and balanced across odd and even runs. The experiment lasted 2 h on average. Experimental timing, audio, and lighting were controlled with Psychtoolbox-3 for Windows (Brainard, 1997).

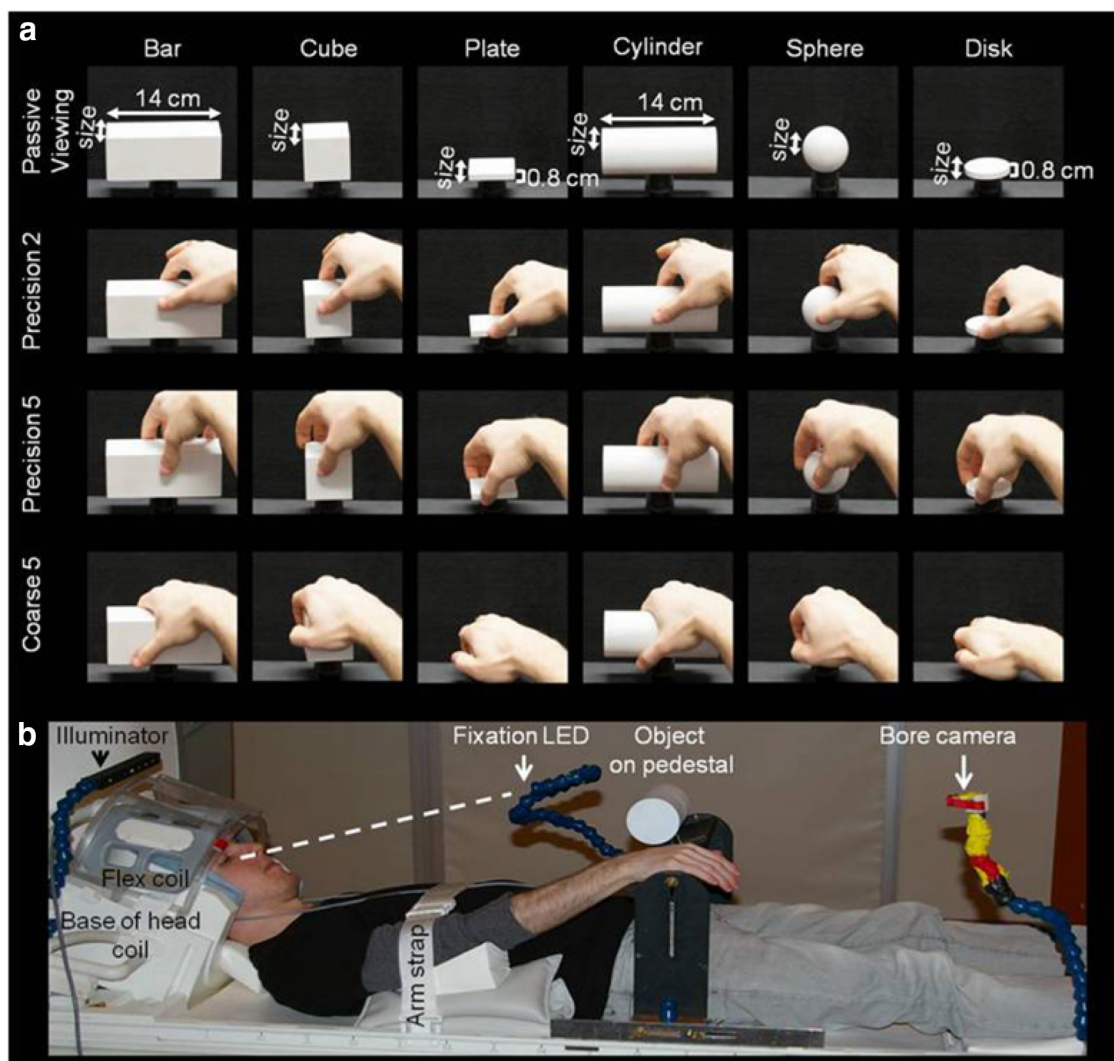


Figure 1. The conditions and experimental setup. *a*, Six objects were shown to participants during passive viewing (first row) and three grasping tasks: Precision 2 (second row), Precision 5 (third row), and Coarse 5 (last row). The six objects were presented in different sizes, varying in the depth of the graspable dimension, which was 3.8, 5.7 (as showed in this figure) or 7.6 cm. Other dimensions were chosen to either retain the object's size (all dimensions of isotropic stimuli, cube and sphere, width of square plate and disk) or to increase or decrease the object's elongation (0.8 cm height for square plate and disk; 14 cm width for bar and cylinder). *b*, The participant's head was tilted to allow fixation of the LED above the object (dotted line), which sat on the pedestal. An LED located behind the flex coil illuminated the workspace during the trial. A bore camera was used to record the events during the trial and to exclude error trials. The participant's right upper arm was restrained with a strap that still permitted movements of the elbow, wrist, and hand. Brain activation was measured with the bottom six channels of a 12-channel head coil, tilted to facilitate direct viewing, and a four-channel flex coil suspended above the forehead.

Data acquisition. All imaging was performed at Robarts Research Institute at the University of Western Ontario (London, Ontario, Canada), using a 3 tesla Siemens TIM Trio MRI scanner.

fMRI volumes were collected using a T2*-weighted, single-shot, gradient-echo echo-planar imaging acquisition sequence [time to repetition (TR) = 2000 ms; slice thickness, 3 mm; in-plane resolution, 3×3 mm; time to echo (TE) = 30 ms; field of view, 240×240 mm; matrix size, 80×80 pixels; flip angle, 90° ; and acceleration factor (integrated parallel acquisition technologies [iPAT] of 2 with generalized, auto-calibrating, partially parallel acquisition reconstruction)]. Each functional volume comprised 34 contiguous (no gap) oblique slices acquired at a $\sim 30^\circ$ caudal tilt with respect to the plane of the anterior and posterior commissure (ACPC), providing near whole-brain coverage.

The T1-weighted anatomical image was collected using a MPRAGE sequence (TR = 2300 ms; TE = 2.98 ms; iPAT, 2; field of view, $192 \times 240 \times 256$ mm; matrix size $192 \times 240 \times 256$ mm; flip angle, 9° ; 1 mm isotropic voxels).

We used a combination of imaging coils to achieve a good signal-to-noise ratio and to enable direct viewing without mirrors or occlusion. Specifically, we tilted ($\sim 20^\circ$) the posterior half of a 12-channel receive-

only head coil (six channels) and suspended a four-channel receive-only flex coil over the anterior-superior part of the head.

Data analysis. Data analysis was performed using Brain Voyager QX 2.8 (Brain Innovation), the BVQX Toolbox (<http://support.brainvoyager.com/available-tools/52-matlab-tools-bvqxtools.html>), NeuroElf (<http://neuroelf.net>), and custom software written in Matlab (The MathWorks).

Preprocessing. Data were preprocessed by applying slice scan-time correction (sinc interpolation for slices acquired in ascending interleaved even-odd order), 3D motion correction (sinc interpolation), and high-pass temporal filtering (cutoff frequency of 3 cycles per run). Functional runs were aligned to the run closest in time to the anatomical volume that was used for functional-to-anatomical coregistration. Both functional and anatomical data were aligned to the ACPC plane and transformed into Talairach space using sinc interpolation. Because the data were analyzed with multivariate pattern analysis (MVPA), no spatial smoothing was applied.

For each participant, we excluded the functional runs where 3D motion parameters indicated abrupt movements > 1 mm translation or 1° in rotation (two runs were excluded in three participants and one run was

excluded in four participants). By off-line video screening, we excluded 3.6% of the trials due to errors in which participants did not correctly perform the task or the experimenter placed the wrong object in the pedestal. We also performed careful quality-assurance inspections to verify that the data were not contaminated by artifacts related to hand movements.

General linear model. Data were analyzed with a random-effects general linear model that included one predictor for each of the 72 conditions convolved with the default Brain Voyager “two-gamma” hemodynamic response function (Friston et al., 1998) and aligned to trial onset. As predictors of no interest, we included the six motion parameters (x , y , and z , for translation and rotation) resulting from the 3D motion correction.

Definition of regions of interest. We investigated a wide range of regions of interest (ROIs) that included early visual areas, motor and somatosensory areas, and areas within the ventral stream (occipitotemporal cortex) and dorsal stream (occipitoparietal and premotor cortex) thought to be specialized for visual recognition versus visually guided actions, respectively (Goodale and Milner, 1992; Culham et al., 2003; Cavina-Pratesi et al., 2007a). Recent studies have suggested that ventral-stream areas may contain more action-related information (Gallivan et al., 2014; Monaco et al., 2014) and dorsal-stream areas may contain more object-related information than previously thought (Bracci and Op de Beeck, 2016).

ROIs were based on the general contrast of all conditions versus baseline, which led to widespread activation related to visual, somatosensory, and motor aspects of the tasks. Initially, we examined random-effects group statistical maps of the general contrast on group-averaged anatomical scans in Talairach space (Talairach and Tournoux, 1988) or on an individual cortical surface. This group statistical map was used for visualization purposes (see Fig. 3) and to determine which regions were commonly activated across the group. That is, foci that were active in individual participants but not apparent in the group analysis were not included as ROIs. Given that right-hemisphere foci may not be reliably identified in sensorimotor tasks performed with the right hand (and that the number of regions included was already considerable), we limited our ROIs to the left hemisphere.

Importantly, ROIs were defined in individual participants based on criteria used in past studies from our laboratory (Gallivan et al., 2011). The ROI approach enabled us to optimally identify regions within individuals (despite variations in stereotaxic location of anatomical landmarks and thus functional regions) without biasing the localization toward any specific hypothesis (Kriegeskorte et al., 2009). Specifically, we localized each ROI based on the activation for each individual in the contrast of all conditions versus baseline (false discover rate, <0.05) using data transformed into Talairach space. Each ROI was localized by selecting the voxels active within a cube (10 mm^3) and ensuring that the cluster contained ≥ 4 functional voxels [$4 \times (3 \times 3 \times 3 \text{ mm}^3) = 108 \text{ mm}^3$]. These cubes were centered at the “hotspot” of activation closest to the expected anatomical landmarks. For example, the aIPS ROI was defined by identifying a cube centered on the hotspot of activation nearest to the junction of the anterior intraparietal and postcentral sulci, as has been done across dozens of studies from our laboratory (see examples of aIPS localization in individual participants in Cavina-Pratesi et al., 2010). Importantly, regions were defined in individuals independently from the subsequent MVPA analysis to avoid any selection bias (Kriegeskorte et al., 2009). Regions were defined based on anatomical and activation criteria rather than stereotaxic coordinates, although Talairach coordinates, as shown in Table 1, were computed based on the average locations of individual ROIs to permit comparisons with other studies.

The following ROIs were identified in the left hemisphere on each individual brain using the following criteria: the putative V1 was identified based on the voxels in, or adjacent to, the calcarine sulcus (Snow et al., 2014); the anterior superior parietal occipital cortex (aSPOC) was localized anterior to the superior end of the parietal occipital sulcus (POS; Culham et al., 2003; Cavina-Pratesi et al., 2010; Monaco et al., 2014); the posterior superior parietal occipital cortex (pSPOC) was localized posterior to the superior end of the POS (Cavina-Pratesi et al., 2010); the LOC was localized inferior to the junction between lateral occipital sulcus of the inferior temporal sulcus, anatomically close to the

Table 1. Talairach coordinates and size of ROIs average across subjects^a

ROI name	Talairach coordinates						ROI size mm ³
	Mean x	Mean y	Mean z	SD x	SD y	SD z	
PMV	−50	−3	34	3	5	3	783
M1	−35	−27	48	4	6	3	888
PMd/FEF	−25	−15	49	3	6	3	853
V1	−6	−81	−2	2	3	6	957
pMTG	−46	−64	−1	2	5	4	800
LOC	−42	−72	−11	3	4	4	728
mIPS	−29	−57	39	3	6	5	754
aIPS	−40	−40	38	6	6	3	843
S1	−44	−32	49	3	8	3	847
pIPS	−27	−60	42	5	5	5	744
aSPOC	−14	−80	29	3	8	7	693
pSPOC	−13	−79	29	4	7	9	703

^aTalairach coordinates (center of mass mean coordinates and SD) and ROI size.

object-selective area LO (Talairach coordinates: $-42, -76, -10$; Bracci et al., 2012); the posterior middle temporal gyrus (pMTG) was localized, in 11 of 12 participants, at the posterior end of the middle temporal gyrus, anatomically close to the hand-and-tool-selective focus (Talairach coordinates: $-47, -66, -2$; Bracci et al., 2012); the posterior intraparietal sulcus (pIPS) was localized at the posterior end of the parietal lobe, close to the anatomical coordinates of the putative caudal intraparietal sulcus (cIPS) area, which has been reported during discrimination of surface orientation in a previous study in humans (Shikata et al., 2008); the PMv was localized at the junction of the inferior frontal sulcus and precentral sulcus (PreCS; Tomassini et al., 2007); the primary motor cortex (M1) was localized at the “hand knob” landmark in the central sulcus (Yousry et al., 1997); the dorsal premotor cortex (PMd/FEF) was localized at the junction of the precentral sulcus and superior frontal sulcus (Tomassini et al., 2007; in absence of specific localizer tasks to distinguish between premotor activity due to eye or hand movements, we called this area PMd/FEF, even though our region is close to the dorsal branch of the superior precentral sulcus, shown to be involved during hand but not eye movements; Amiez et al., 2006); the middle intraparietal sulcus (mIPS) was localized halfway up the length of the IPS, on the medial bank (Gallivan et al., 2011); the aIPS was localized at the junction of the IPS and postcentral sulcus (PostCS) (Culham et al., 2003); the somatosensory cortex (S1) was localized by selecting the voxels in the postcentral gyrus and sulcus (Gallivan et al., 2011).

RSA. Multivoxel pattern analysis was performed using the method of RSA to measure the nature of coding in each region (Kriegeskorte et al., 2008; Nili et al., 2014). Essentially, by examining in an ROI the degree to which each of the 72 conditions evoked a similar pattern of brain activity to each of the others, the nature of the neural code (or representational geometry; Kriegeskorte and Kievit, 2013) can be assessed. We then compared the measured neural codes to three hypothetical motor and three hypothetical visual codes.

At each voxel, baseline z -normalized β estimates of the blood-oxygen-level-dependent response, representing the mean-centered signal for each voxel and condition relative to the SD of signal fluctuations, were extracted for each of the 72 conditions at every functional voxel in each ROI. This procedure was repeated for all runs. Error trials were excluded in the matrix of β weights.

First, we computed representational similarity matrices (RSMs) for each ROI. Using Pearson correlations of voxel β weights across voxels within the ROI, these RSMs measure how similar the pattern of activity evoked by each condition is to the pattern of activity evoked by every other condition. To test the within-subject reliability of these β weights, we used the cross-validation procedure of splitting the data in two halves (odd and even runs). For each set of runs, the β weights were averaged separately for each condition and voxel. For each voxel, the average β weight was normalized to a mean of zero by subtracting the mean β value across conditions from the β value of each condition in that voxel (Haxby et al., 2001). Pearson correlations (r) were computed between the pattern of β 's across voxels in each condition during odd and even runs, resulting in a 72×72 asymmetrical RSM, and then Fisher transformed to yield a

similarity metric with a Gaussian distribution. A second, nonsplit symmetrical RSM, obtained correlating conditions averaged across all runs, was calculated to visualize the similarity structure of the activity patterns for all 72 conditions (or separately for the 54 conditions during grasping and the 18 conditions during passive viewing) in each ROI using multidimensional scaling (MDS) transformation. To do so, for each ROI the nonsplit RSM was averaged across participants and subjected to MDS (using the Matlab function `mdscale`, which performs nonmetric MDS for two dimensions with the squared stress criterion) to generate 2D representational plots. To enable direct comparison between regions, all the MDS plots were created using the same scale for all regions. While nonsplit RSMs, which have $r = 1$ along the diagonal because the correlations are based on the same data, are necessary for MDS plots shown in Figures 3 and 4 (because a given condition must be zero distance from itself), the split RSMs do not artificially inflate the similarity of same-condition correlations relative to between-condition correlations and were used in the following steps of the analyses. To test whether each region contained information about condition differences, we measured correlations between the RSM in each region and a model that contrasted diagonal versus off-diagonal cells. This provides a test as to whether voxel patterns show higher similarity between identical conditions than between different conditions. Because correlations in all regions were significant ($p < 0.05$), we proceeded with testing specific models as described below.

Second, we tested how well the representational data in each ROI could be explained by models based on the similarity between motor attributes (grasping tasks) and/or visual attributes (grasping and passive-viewing tasks). Although the investigation of grip type classically includes Coarse 5 (also called Power grip, when applied to elongated objects, or Whole-hand grasp) and Precision 2 (for review, see Castiello, 2005), here we also included a third type, Precision 5, in which participants used all five digits, but not the palm, to grasp an object precisely. This enabled us to test whether the relevant dimension was the number of digits involved (Digit Model), or the amount of precision required (Precision Model). Neurophysiological studies have found neurons selective for elongated and flat objects in the macaque cIPS (Sakata et al., 1998), which provides input to the AIP and which may have a human homolog (Shikata et al., 2008). Moreover, elongation may influence human sensorimotor processing (Fang and He, 2005; Sakuraba et al., 2012). To test whether these 3D properties were coded in the human brain, we created the Elongation Model contrasting flat, isotropic, and elongated objects. We also created the Shape Model to test whether the geometrical shape (rectangular or square) of the object was represented during grasping and/or during passive viewing. Finally, we created the Size Model to test the representation of the object's size along the grasped dimension, shown to be relevant during grasping (Ganel and Goodale, 2003). These models are not exhaustive for all possible dimensions in the stimuli. Indeed, some dimensions can be derived by the combination of the aforementioned models; for example, the aspect ratio of the stimuli is a combination of object elongation and size, as the bar and cylinder have a different aspect ratio depending on size, while aspect ratio remains constant for a disk, plate, sphere, and cube of different sizes. We also did not explicitly test the effect of the object's volume (i.e., which would determine its weight) because our participants did not lift the objects and because the effects of this variable have been carefully examined previously (Gallivan et al., 2014).

Specifically, in the RSMs for grasping (54×54) and passive viewing (18×18), we tested five models (see Fig. 5). Motor models during grasping tasks were as follows: (1) Precision Model: regardless of object properties, the two grips that require precision (Precision 2 and Precision 5) are represented similarly to each other but differently from the less precise Coarse 5 grip; (2) Digit Model: regardless of object properties, the two grips that use five digits (Precision 5 and Coarse 5) are represented similarly to each other but differently from the grip that uses two digits (Precision 2). Visual models (tested separately during grasping and during passive viewing) were as follows: (1) Shape Model: regardless of grip and other visual properties, square objects (plates, cubes, and bars of all sizes) are represented differently from round objects (disks, spheres, and cylinders of all sizes); (2) Size Model: regardless of grip and other visual properties, objects are represented similarly based on the size of the

dimension to grasp; small, medium, and large objects are differently represented; (3) Elongation Model: regardless of grip and other visual properties, objects are represented similarly based on their elongation [e.g., flat objects (square plates and round disks) are represented similarly to each other but differently from isotropic objects (cubes and spheres) and elongated objects (cylinders and bars)].

To assess each code, we computed the Spearman correlation between the individual Fisher-transformed split RSMs and each model and plotted the results for all regions during grasping and passive viewing. We then used t tests to determine whether these correlations were significantly >0 within each ROI and corrected the p values for the number of models tested within a region [during grasping, $p < 0.010$ (0.5/5 models); during passive viewing $p < 0.016$ (0.05/3 models)]. The results are shown with markers in Figures 6 and 7. Next, to compare the strength of significant models in a region, we computed pairwise t tests (noncorrected $p < 0.05$) between all models within each region and showed the models that provided a significantly better fit of the data with filled markers in Figures 6 and 7. To estimate the maximum correlation value expected in a region, given the variability across subjects for each region, we computed the noise ceiling by using a leave-one-out approach correlating each subject's RSM with the average RSM of the other subjects' RSMs (for a similar approach, see Bracci and Op de Beeck, 2016). The resulting correlation values are shown with the dotted line in Figures 6 and 7. Since the noise ceiling indicates how consistent the RSMs are across participants, a high noise ceiling indicates highly consistent patterns of correlations within that region across participants, while very low values of the noise ceiling indicate high variability in the RSM between participants within that region. Therefore, the performance of the models should be interpreted with caution in a region where the noise ceiling is low because in that region data are not very consistent between participants.

Third, we used two techniques to visualize the similarity of representations between the 12 ROIs. Because of the possibility that regional similarities differed between grasping and passive-viewing tasks, they were examined separately (on a 54×54 condition matrix for grasping and an 18×18 condition matrix for passive viewing). Spearman correlations were computed between the split odd–even RSM of each region and every other region, producing a symmetrical 12×12 RSM for the grasping task and another one for the passive-viewing task. MDS was applied to these 12×12 RSMs to produce a 2D plot showing the similarities between the representations in the 12 ROIs during grasping (see Fig. 8*b*) and passive viewing (see Fig. 9*b*). These RSMs were also used to create hierarchical cluster plots (using the weighted distance criterion of the Matlab function `linkage`).

Searchlight analysis. To test whether our ROI analyses may have missed additional areas that code the visual and motor attributes we examined, we performed, with custom-made Matlab scripts, a whole-brain searchlight RSA, using a 15-mm-diameter sphere on each individual brain (Kriegeskorte et al., 2006). For each model (see Fig. 5), a second-level analysis was performed to generate t maps showing voxels where the correlations between the individual RSM and the model were statistically >0 . The resulting t maps were thresholded at $t > 3.4$ ($p < 0.005$) with a cluster correction using a Monte Carlo simulation method (Forman et al., 1995) performed with the function `AlphaSim` in `NeuroElf` (<http://neuroelf.net/>).

Results

Representational similarity between all 72 conditions within each ROI

Figure 2 shows the 72×72 RSMs for each of the 12 ROIs. Each point in the matrix represents the similarity of the neural code evoked by one condition during the odd runs with that evoked by another condition during the even runs. Warmer colors indicate the conditions that evoked more similar patterns of activity in the ROI. Structure in the matrix reveals the representational geometry of the neural code. For example, the fine checkerboard pattern apparent in the V1 RSM data indicates strong similarity (red squares) between bars, cubes, cylinders, and spheres (nonflat objects), when these objects were grasped with the similar degree of

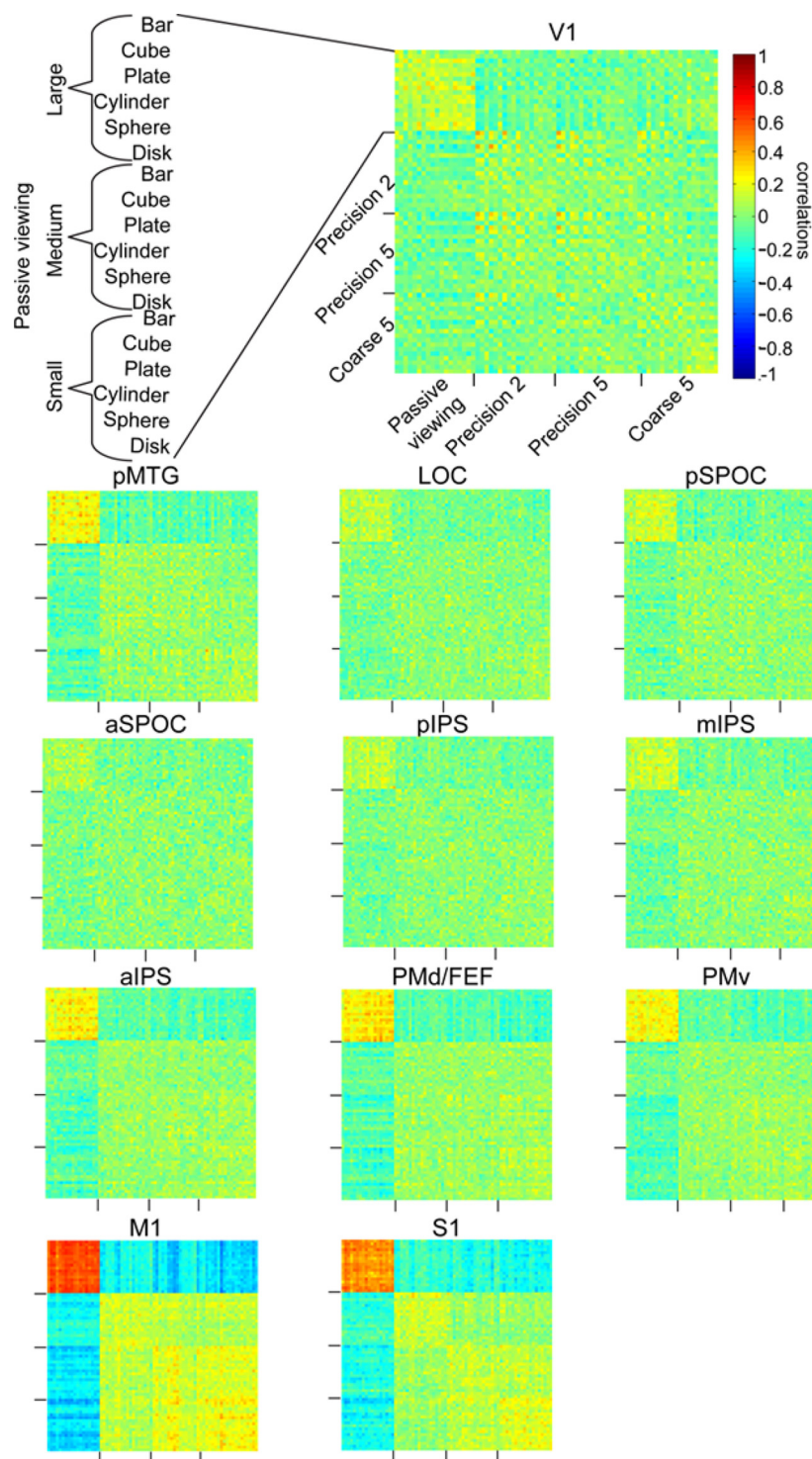


Figure 2. RSMs for each region. Each point in the matrix indicates the similarity of the patterns of brain activity evoked by a given pair of conditions. Conditions are ordered (up–down and left–right) by task (passive viewing, Precision 2, Precision 5, and Coarse 5), with task divisions indicated by ticks at the border of the matrices. Within each task, 18 objects are presented in the specific order and size, as shown in the enlarged area for passive viewing in V1. The color within each of the 72×72 cells indicates the strength of the correlation among spatial patterns in the region during odd and even runs. The data were split in odd and even runs, and thus noise may make the matrix asymmetric.

precision (Precision 2 and Precision 5). Note that this pattern of similarity is also shown in the MDS plot for V1 in Figure 3, where green icons for bars, cubes, cylinders, and spheres (corresponding to conditions where those objects were grasped using Precision 2) overlap with orange icons for the same objects

(corresponding to conditions where objects were grasped using Precision 5). Visual inspection of all the RSMs reveals that the strongest similarity in most of the regions is between conditions within the same task modality (i.e., within passive viewing and within the three grasping tasks). Conditions within different task modalities (i.e., passive viewing vs grasping) share very low similarity, particularly in motor (M1), somatosensory (S1), and high-level sensorimotor (PMd/v and aIPS) regions. Although the RSMs do not show a striking difference between the diagonal cells compared with off-diagonal cells, two considerations are important. First, RSMs are often generated using nonsplit data, where correlation values along the diagonal must be 1 (because the data are fully correlated with the same data); however, here we preferred to show the split data (with correlations between even and odd runs) because they do not artificially exaggerate the diagonal correlations and thus provide a better estimate of the maximum correlation that could be expected. Second, as our data will show, most regions do not exclusively represent only one combination of object and task (as would be found along the diagonal) but rather a combination of attributes. To visualize the major aspects of the organization of these condition groupings, 2D MDS plots of representational spaces are shown for each of the 12 ROIs in Figure 3. Although the group data are shown on the inflated surface of one individual's left hemisphere, the analyses were based on ROIs extracted separately for each individual. The distribution of the icons in an MDS plot represents the neuronal similarity between conditions in the specific region: the closer together the icons are, the more those conditions are similarly represented in the specific brain area. By contrast, the farther apart the icons are, the more their neuronal representation is distinct. This way of plotting the correlations between the experimental conditions allows visualization of similarities and differences among our 72 conditions in a data-driven approach, since the MDS plot groups the data without any assumption about categorical organization. It is worth emphasizing that these MDS plots qualitatively show only the two dimensions that account for the most variance in the data. Although the inclusion of additional dimensions would account for more of the variance, they are also more difficult to visualize and interpret (as our own inspection of 3D plots revealed). Importantly, our quantification and statistical testing of various models (presented below) relies on the full (N -dimensional) data.

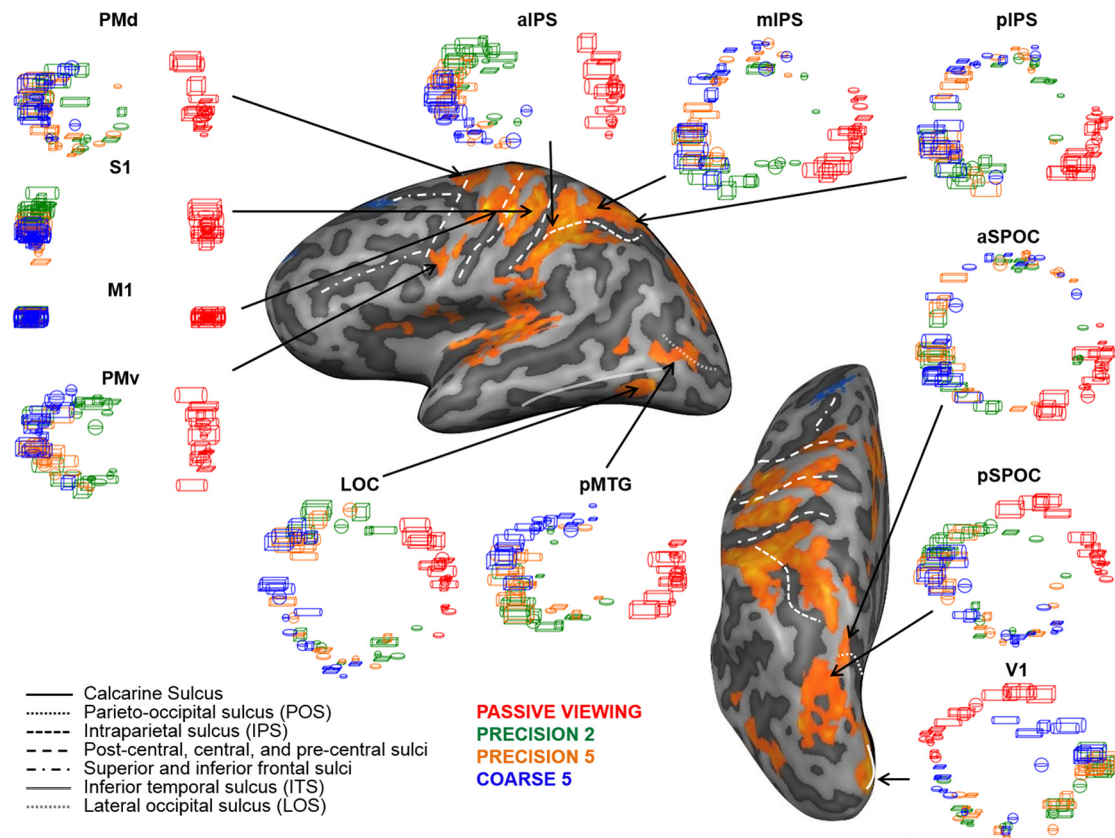


Figure 3. The random-effects contrast among all conditions and baseline projected onto the inflated cortical surface (sulci, dark gray; gyri, light gray) of the left hemisphere of a representative participant, thresholded at false discovery rate-corrected $p < 0.05$. Lines on the inflated brain localize key sulci. From each region, the MDS plots indicate the similarities of the representations evoked in 2D space by the 18 objects (indicated by icons) during passive viewing (red) and grasping in the Precision 2 (green), Precision 5 (orange), and Coarse 5 (blue) conditions. Across regions, the spatial organization of the icons shows a clear separation of the red icons from the other icons, indicating that the neuronal representations differ between passive-viewing and grasping conditions. This grouping based on task is particularly strong in M1, where it swamps any differences based on other dimensions of the data, and weaker in V1, where data representations are also affected by object elongation.

Visual inspection of MDS plots reveals that data are organized based on object dimensions (such as clustering based on object elongation) and task types (as indicated by color groupings). This organization varies across regions. The MDS plot in M1 reveals a strong distinction between passive viewing and the three grasping tasks. The MDS plot for S1 is similar, but with some added distinctions among the three grasping tasks. Pronounced differences among the grasping tasks and passive viewing are also seen in premotor cortices (PMd and PMv), anterior areas within the posterior parietal cortex (aIPS, mIPS, pIPS) and the posterior middle temporal gyrus, and to a lesser degree in SPOCs (aSPOC and pSPOC), the LOC, and V1. Within V1, in addition to task differences, effects of object elongation are also apparent. Elongation effects are also apparent in other regions. The spatial organization of object properties in relation to the specific task differs among regions, indicating different sensitivity for different stimulus dimensions across regions. To better visualize the effect of object properties within each task modality, we created MDS plots separately for grasping and passive viewing. Figure 4 shows that during grasping, motor areas like M1, S1, and the PMv represented conditions based on the type of action, as indicated by the grouping of icons based mainly on their colors. In contrast, the MDS plots in the remaining regions revealed grouping based on both object properties and type of action. During passive viewing, groupings were less clear as the icons were generally more scattered, with the exception of

M1 and S1 where there is no clear distinction between icons. In the remaining regions, elongated objects tend to group separately from the others.

Representation of visual and motor information in each ROI separately during grasping and passive viewing

To quantify the degree to which the spatial organization shown in Figure 4 reflected the representation of object properties, of the specific task, or both, we tested the extent to which specific theoretical models correlate with the neural data. In particular, we computed correlations between the split RSMs for each participant and three visual models, where object elongation, size, and shape were represented regardless of the task, and three motor models, where the degree of precision and the number of digits used to grasp the object were represented independently of object properties (Fig. 5; see Materials and Methods). Because the MDS plots in Figures 3 and 4 show that each region represents information differently during grasping versus passive viewing, we analyzed the data first for the three grasping tasks combined (56×56 matrix) and later for the passive-viewing condition (18×18 matrix) alone. For the grasping data, we tested both motor models (Precision and Digit models) and all three visual models (Shape, Size, and Elongation models); for the passive-viewing data, only the visual models were tested.

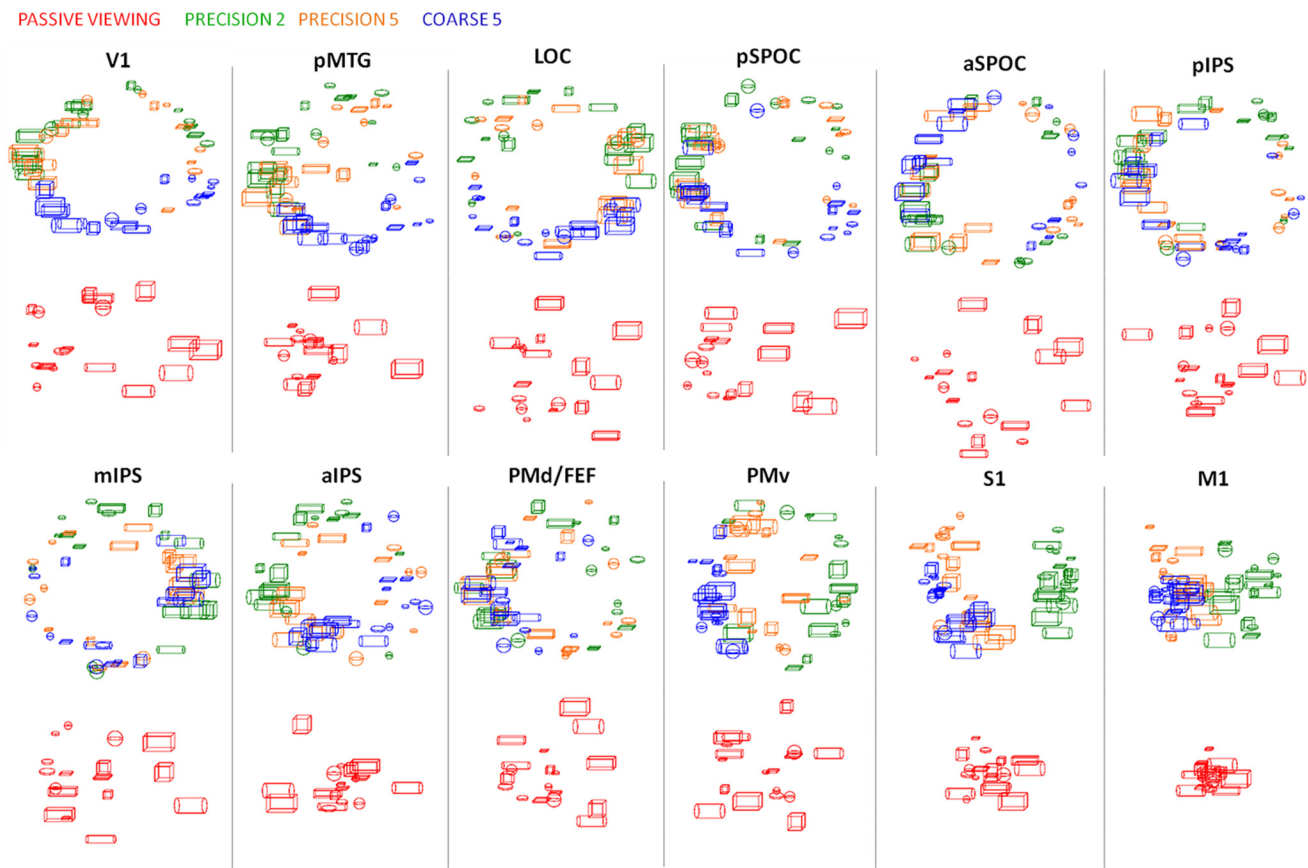


Figure 4. From each region, the MDS plots indicate the similarities of the representations evoked in 2D space by the 18 objects (indicated by icons) separately during grasping (first row: Precision 2, green; Precision 5, orange; Coarse 5, blue) and passive viewing (second row: red). Note that in MDS plots, the orientation of the axes can vary; that is, similar attributes may be coded in an area during grasping and passive viewing, but the plots may have different layouts. For example, the aSPOC shows a grouping by elongation during both tasks, but the layout is opposite (elongated objects appear on the left for grasping and on the right for passive viewing).

Grasping tasks

As expected from our factorial design, the patterns within RSMs from many areas could be accounted for by multiple models. Indeed, in most regions (except the aSPOC), >1 model performed significantly better than chance ($r = 0$, $p < 0.01$), as indicated by round markers in Figure 6 (Table 2). Filled markers in Figure 6 show models that provided a significantly better fit of the data than other models (based on pairwise comparisons between all models within each ROI; $p < 0.05$).

Across regions, we found that Elongation and Digit models explain the data better than the Shape, Size, and Precision models. The strength of each model's coding was compared with the noise ceiling, representing an estimate of the maximal coding strength observable, given the degree of noise in the data, the lower bound on the noise ceiling (Nili et al., 2014; see Materials and Methods). In general, the Elongation and the Digit models were the only models that approached or exceeded this noise ceiling, suggesting that these models perform about as well as could be expected given the variability in the data across participants. Different levels of the noise ceiling across regions indicate that the neural representation is more consistent across participants in primary visual, motor, and somatosensory areas than in the remaining regions. Given these different levels of noise ceiling among regions, we statistically compared models within but not among regions.

Areas demonstrated a gradient from visual to visuomotor to motor coding, with only visual and only motor sensitivity in early

visual and motor areas at the extremes of the gradient, and both visual and motor sensitivity in ventral-stream and dorsal-stream areas. Specifically, in V1, as well as the aSPOC and the pSPOC, the visual Elongation model explained the data significantly better than all the other visual and motor models. By contrast, in M1, S1, and the PMv, the motor Digit Model was significantly stronger than all the other motor and visual models. The Digit Model as well as Shape and Elongation showed comparable significance in the aIPS. This visuomotor sensitivity in the aIPS and exclusive motor sensitivity in the PMv confirmed our predictions, suggesting that the aIPS combines object properties into the specific hand configuration, while the PMv codes motor aspects of the grasping task. In line with our last hypothesis, we found similar sensitivity in ventral-stream and dorsal-stream regions during grasping, showing coding of both object elongation and number of digits in the pMTG, the LOC, the pIPS, the mIPS, and the PMd. Beyond object elongation and the number of digits, the PMd and the pMTG also coded object size, while the pMTG also coded the amount of precision used to grasp the object.

Passive-viewing task

Although we were primarily interested in analyzing the factors contributing to neural representations during grasping, the passive-viewing condition provides a valuable way to measure whether visual information is differently processed when no action is required. Notably, given that the vast majority of neuroimaging studies of object recognition have used 2D images (or

occasionally 3D simulations), here we can investigate the coding of real objects in the human brain, as has been done previously in the macaque (Schaffelhofer et al., 2015). There are both theoretical and empirical reasons to expect that the representations of real objects may differ from those of visually matched images. Theoretically, real objects are visually richer (including for example multiple and consistent cues to 3D structure) and afford genuine actions (and adults never attempt to grasp objects in images, although infants may; DeLoache et al., 1998). Empirically, recent fMRI data suggest that both the ventral and dorsal streams process real objects differently from visually matched photos. Specifically, while passive viewing of 2D photos elicits adaptation in ventral and dorsal streams, the effect was weak or absent for real objects (Snow et al., 2011). Critically, ongoing work from our laboratory finds substantial differences in the representations (based on RSA) of real objects compared with visually matched photos (Snow et al., 2015).

Thus to examine the representations of real objects, we measured the similarity between the visual models (Fig. 5) and the individual RSMs in each region during passive-viewing conditions. The resulting correlation averages across participants are shown in Figure 7, where models that performed significantly better than chance are shown with round empty markers, as in Figure 6 (Table 3). Notably, the absence of filled markers in Figure 6 indicates no statistical difference between models.

Overall the representation of object properties during passive viewing followed a similar gradient as during grasping, with visual properties of the object represented exclusively in visual and parietal areas, except for object elongation in M1. Interestingly, during passive viewing the representation of visual information changed compared with the reported sensitivity during grasping. Indeed, visual areas represented object elongation and shape at the same strength in V1, the pMTG, and the pSPOC. Thus, visual areas no longer showed a strong preference for object elongation. By contrast, the LOC and the aSPOC coded object shape only, while the mIPS and M1 coded only object elongation. The LOC, aSPOC, pIPS, aIPS, PMd/FEF, and S1 lost their sensitivity for object elongation observed during grasping. The fact that most of the models did not exceed the noise ceiling indicated that the neural representation of the visual models during passive viewing was variable across participants. The noise ceiling around zero in the pMTG and M1 indicated that the amount of information contained in these regions is highly variable across participants, reaching chance level if some participants are excluded.

Representational similarities across regions

To more formally evaluate similarities in the neural code across different regions, we computed the structure of representational similarity among ROIs. We used their RSMs to create a hierarchical cluster plot and a 2D MDS plot. Given the observed differ-

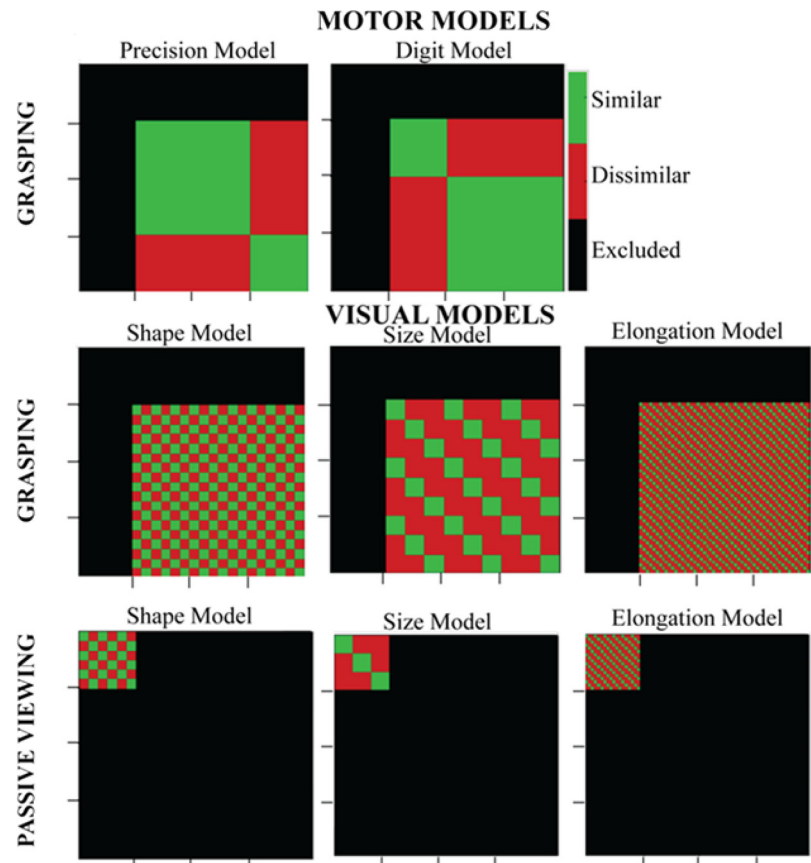


Figure 5. The motor and visual models tested. For each model, each of the 72×72 cells (ordered as in Fig. 2) is color coded to show which conditions are expected to be most similar (green) or dissimilar (red). Each model was correlated with the RSM for each area (Fig. 2) to derive the graphs in Figures 5 and 6. For grasping trials (lower right 56×56 region of the matrix), both motor and visual models were tested. For passive-viewing trials (upper right 18×18 region of the matrix), only visual models were tested.

ences in processing pathways for grasping and passive-viewing tasks, we analyzed these tasks separately.

Grasping task

Both panels in Figure 8 show that similarities among brain regions form two main groups during grasping: motor areas S1, M1, PMd/FEF, and PMv group together and separately from the remaining regions. The second cluster contains early visual and ventral areas during grasping. Interestingly, dorsal-stream regions pIPS, mIPS, and aIPS are also part of this latter cluster, suggesting that these regions share information with visual areas. This gradient of information can be visualized along the x-axes of the MDS plot where areas are organized along the aforementioned visuo-to-motor gradient. While the main findings can be observed in both plots (e.g., S1, M1, PMd, and PMv are part of the same group based on their neuronal similarity), small inconsistencies between the two plots (e.g., S1 and M1 share a lot of similarity in the hierarchical plots, and less in the MDS plot) can be related to different assumptions: while the hierarchical plots assume the presence of a structure in the data, the MDS plot does not rely on this assumption (Nili et al., 2014). The clusters showed in Figure 8 is in line with the significant models shown in Figure 5, indicating that the cluster formed by S1, M1, and the PMv could reflect the exclusively motor sensitivity of these regions for the motor aspects of the grasping task, in particular for the number of digits and the grip type. The cluster formed by the remaining regions might reflect the common sensitivity for visual information, which, depending on the extent of this common

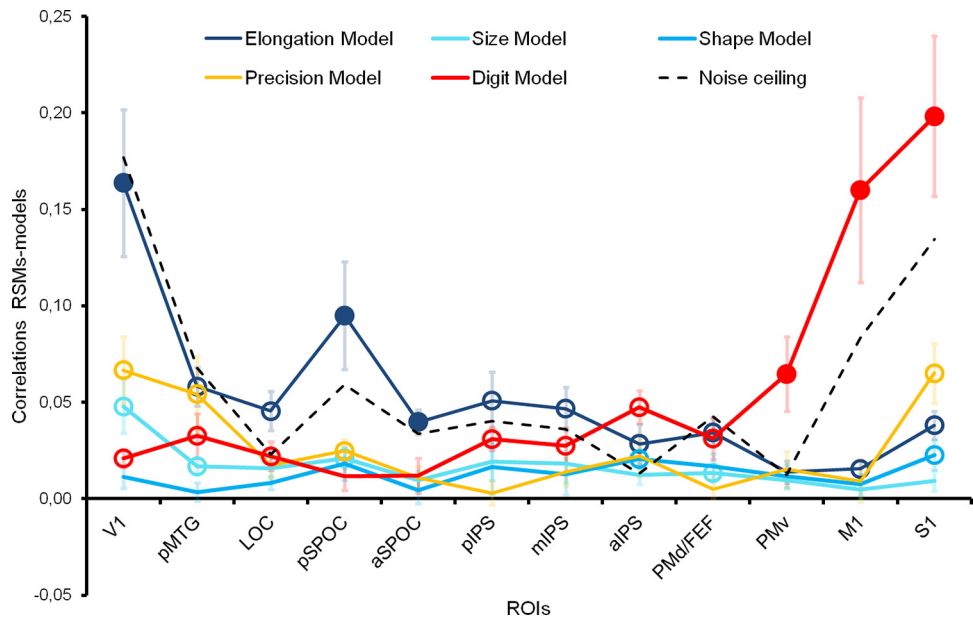


Figure 6. Average correlations between individual RSMs and the visual (cold colors) and motor (warm colors) models across ROIs during grasping tasks. Round markers indicate models significantly different from zero (p corrected for the number of models; $p < 0.010$). The strongest models, resulting from *post hoc* pairwise comparisons within each ROI ($p < 0.05$) are shown with filled markers. A single filled marker in an ROI indicates a winner-take-all patterns, whereas multiple empty markers indicate multiple neural codes of a similar strength. The black dotted line shows the lower bound on the noise ceiling—an estimate of the maximal coding strength expected given the measured noise in the data. Error bars show the SE.

Table 2. p values resulting from t test versus zero during grasping^a

Model name	ROI name											
	V1	pMTG	LOC	pSPOC	aSPOC	pIPS	mIPS	aIPS	PMd/FEF	PMv	M1	S1
Digit Model	0.0006 ^b	0.0088 ^b	0.0079 ^b	0.0730	0.1105	0.0002 ^b	0.0066 ^b	0.0001 ^b	0.0080 ^b	0.0033 ^b	0.0033 ^b	0.0004 ^b
Precision Model	0.0017 ^b	0.0095 ^b	0.0167	0.0010 ^b	0.0572	0.3171	0.0368	0.0337	0.1598	0.0625	0.2100	0.0010 ^b
Shape Model	0.0478	0.2345	0.0196	0.0343	0.2736	0.1419	0.1280	0.0007 ^b	0.0140	0.0368	0.0936	0.0075 ^b
Size Model	0.0037 ^b	0.0020 ^b	0.0418	0.0013 ^b	0.0446	0.0344	0.0413	0.0142	0.0070 ^b	0.0277	0.1630	0.0550
Elongation Model	0.0009 ^b	0.0001 ^b	0.0007 ^b	0.0038 ^b	0.0000 ^b	0.0027 ^b	0.0007 ^b	0.0087 ^b	0.0004 ^b	0.0177	0.0025 ^b	0.0002 ^b

^aResults of t test for each model in each ROI during grasping.

^bSignificant p values (p corrected for the number of models; $p < 0.010$).

sensitivity, creates the subgroups of primarily visual areas (V1, pSPOC, aSPOC, LOC, and pMTG) and both visual and motor areas (aIPS, mIPS, pIPS).

Passive-viewing task

Figure 9 shows that neural similarities among brain regions form two main groups during passive viewing: motor areas S1, M1, and PMv group together and separately from the remaining regions. The second cluster contains early visual and ventral areas. Similar to results during grasping, dorsal-stream regions pIPS, mIPS, aIPS, and PMd/FEF are part of this latter cluster, suggesting that these regions shared similar information with visual areas during passive viewing.

Because most of the regions that group together in the cluster analysis, like S1 and M1, are also anatomically close, one could argue that Figures 8 and 9 reflect not representational similarity but physical proximity. We do not think this arises from artifacts related to physical proximity (for example if noise followed a gradient) as these would not be expected to produce the intuitive findings here (for example, the dominance of visual models in visual areas and motor models in motor areas) nor the different inter-regional similarities found for grasping and passive viewing. Rather, the effect of physical distance likely arises from a fundamental principle of brain organization: brain regions with similar function evolve and develop to be physically close to min-

imize the amount of “wiring” (white matter fibers) that connects them (Van Essen, 1997). Such patterns are often observed in anatomical and functional connectivity studies (Young, 1992) and this is found across diverse methods from tract tracing in nonhuman primates to resting-state connectivity in humans. By such an account, representational similarities are better explained by connectivity than physical proximity per se (although connectivity and proximity are closely related). For example, in our data, M1 and S1 group separately from the other areas, which is consistent with their position in the sensory-motor hierarchy, rather than appearing “between” parietal and frontal areas, as they are in terms of physical location.

Searchlight analysis

To test whether the ROI analysis may have missed regions that code the visual and motor attributes we examined, we ran a whole-brain searchlight RSA during grasping and passive viewing. Since Precision and Shape models did not survive cluster correction, only the t maps of the remaining models are shown in Figure 10.

Although the searchlight (Fig. 10a) identified similar coding during grasping as the ROI analysis, there were also some differences. Both approaches showed that the predominant visual model was Elongation, though the searchlight revealed that size coding also accounted for data in early visual areas. In fact, the

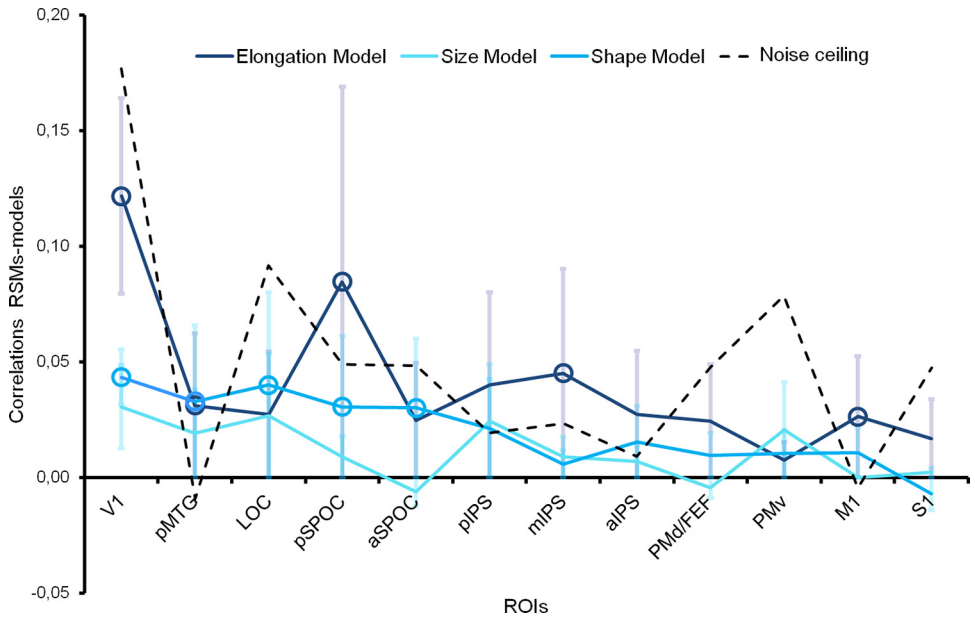


Figure 7. Average correlations between individual RSMs and the visual models across ROIs during passive viewing. As in Figure 5, markers indicate models that significantly differ from chance (p corrected for the number of models; $p < 0.016$). As in Figure 5, filled markers would indicate models that have stronger coding than others within that ROI. An absence of filled markers indicates no statistical difference among significant models. The black dotted line shows the lower bound on the noise ceiling—an estimate of the maximal coding strength expected given the measured noise in the data. Error bars show the SE.

Table 3. p values resulting from t test versus zero during passive viewing^a

	ROI name											
Model name	V1	pMTG	LOC	pSPOC	aSPOC	pIPS	mIPS	aIPS	PMd/FEF	PMv	M1	S1
Shape Model	0.0025 ^b	0.0137 ^b	0.0011 ^b	0.0056 ^b	0.0147 ^b	0.0418	0.2970	0.1795	0.2065	0.2904	0.1291	0.7578
Size Model	0.0672	0.1380	0.0181	0.3167	0.6809	0.0593	0.2724	0.3451	0.5985	0.1564	0.5013	0.3789
Elongation Model	0.0093 ^b	0.0151 ^b	0.0403	0.0041 ^b	0.0467	0.0617	0.0096 ^b	0.0392	0.1289	0.3390	0.0080 ^b	0.1113

^aResults of t test for each model in each ROI during passive viewing.

^bSignificant p values (p corrected for the number of models; $p < 0.016$).

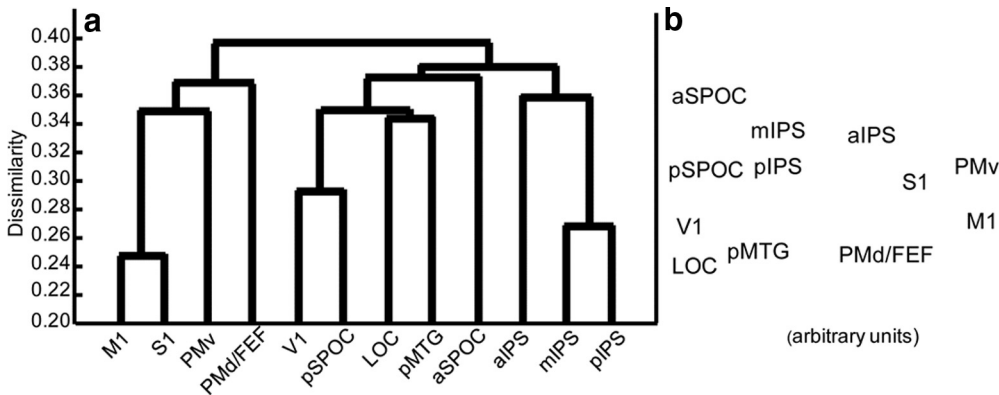


Figure 8. Visualization of similarity of regions in their neural codes during grasping. **a**, Hierarchical clustering, which shows the hierarchically grouped regions with the most similar representations, indicated two broad clusters: one comprising motor–premotor areas and another comprising visual–parietal areas. **b**, A 2D MDS plot also showed two main groups, as seen in panel **a** (left–right axis), and additionally showed that parietal representations fell between visual and motor areas.

searchlight revealed that elongation and size coding extended beyond the putative-V1 (calcarine) and ventral-stream foci selected for the ROI approach, to include the cuneus (where the V1/V2/V3 process the lower visual field where our stimuli were presented) and the superior/lateral occipital cortex (which includes midlevel visual areas, such as V3A). As with the ROI approach, the searchlight revealed that the Elongation Model predominated throughout the IPS and in the motor cortex.

Moreover both approaches showed that the Digit Model was more influential than the Precision Model [including in areas not examined in the ROI analysis, such as the somatosensory area (S2)] in the motor, somatosensory, premotor, and aIPS regions, as in the ROI analysis, but also in the supplementary motor area and regions surrounding the insula, including S2. Although an effect of Digit and Elongation models appears on the superior temporal gyrus and could include the auditory cortex, we suspect

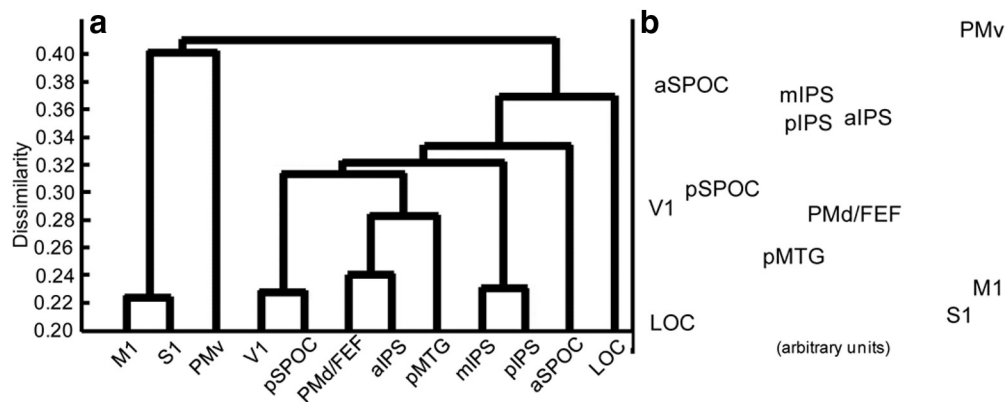


Figure 9. Visualization of similarity of regions in their neural codes during passive viewing. *a*, A hierarchical clustering again indicated two broad clusters, one that includes M1, S1, and the PMv and one that includes the PMd and visual-parietal areas. *b*, A 2D MDS plot also indicated two main groups seen in panel *a* (left–right axis), and additionally showed that parietal representations fell between visual and motor areas, while the PMv is separated from other regions.

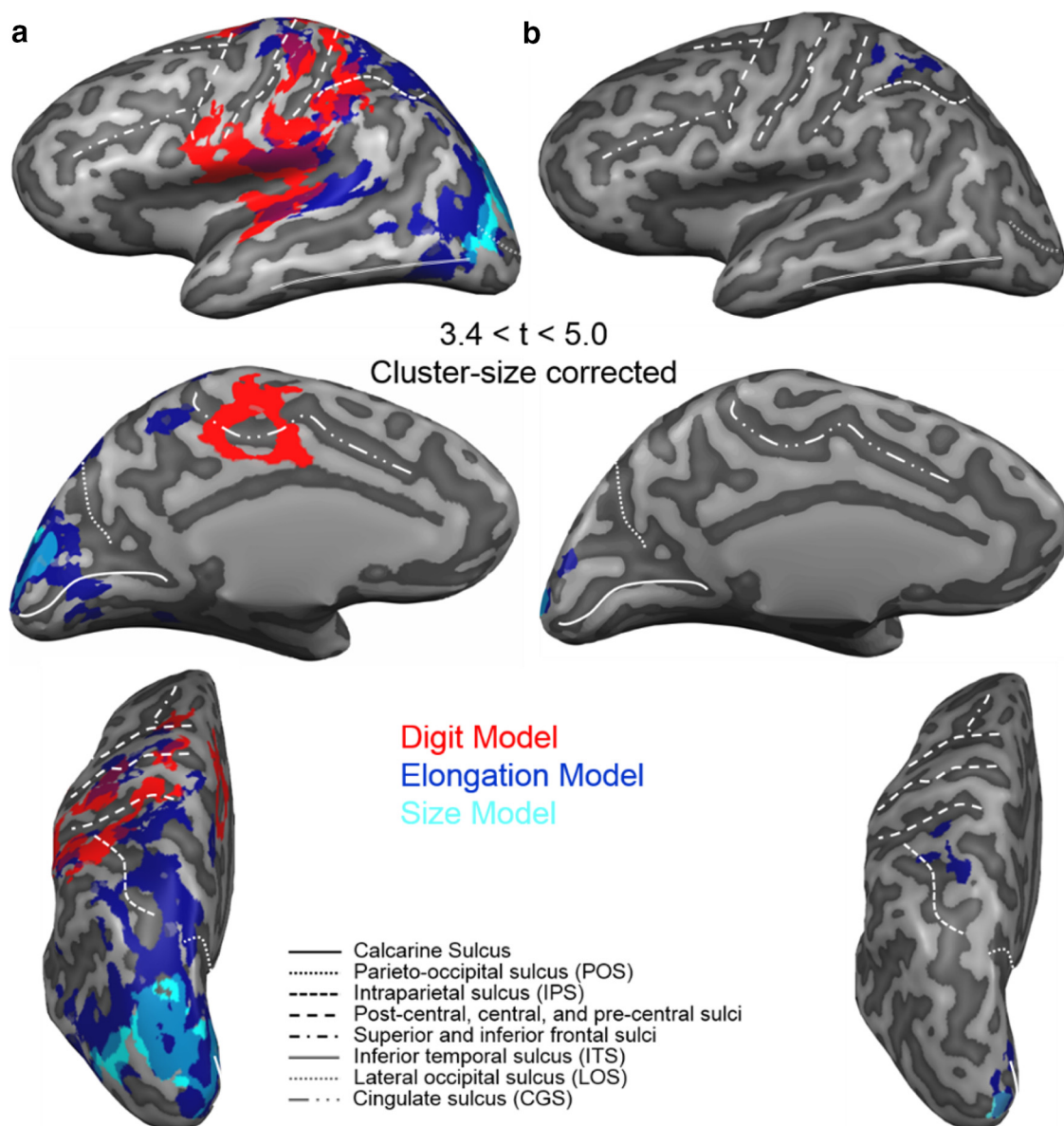


Figure 10. Results from the searchlight RSA. *a*, Group *t* maps showing significant motor and visual models during grasping. *b*, Group *t* maps showing significant visual models during passive viewing. As in Figure 3, group results are rendered on the cortical surface of one individual participant.

this is an artifact of the projection of volumetric group data on a cortical surface (such that insula S2 activation, which has a wide spread based on the searchlight diameter, appears projected on both banks of the Sylvian fissure). Alternatively, it may be that the auditory cortex can decode the audio cues related to Digits; however, there were no such cue differences for Elongation conditions, making attribution to S2 more plausible.

For passive viewing (Fig. 10*b*), the searchlight confirmed the weak coding of object properties within regions identified in the ROI analysis. During passive viewing, one might have expected that the searchlight analysis would have revealed coding of object properties in ventral-stream regions like the medial fusiform gyrus, shown to code tool manipulability (Mahon et al., 2007). Unlike Mahon et al., who used pictures of familiar objects, we used real unfamiliar objects, so that their manipulability was not based on prior experience.

In some cases, the searchlight found regions missed by the ROI analysis, which is the benefit of whole-brain approaches over ROI analysis. In other cases, the ROI analysis revealed similarities between the model and data that were missed by the searchlight. This is to be expected as the searchlight analysis is less powerful (more conservative) than the ROI analysis for two reasons: (1) because of the required correction for multiple comparisons necessary when considering voxels across the whole brain and (2) because it is vulnerable to variations in the stereotaxic location of functional areas across subjects.

To summarize, this study revealed that visuomotor sensitivity across regions involved in grasping is mainly driven by the representation of object elongation above object shape and size in visual areas, and by the number of digits above grip type in motor areas. Within this continuum, ventral and parietal regions represented visual and motor properties at a similar strength during grasping. This reported sensitivity to object elongation is specific to the grasping task as during passive viewing the representations of object shape and/or size were represented together with object elongation, indicating that the visual information about the object is influenced by task demands in these regions.

Discussion

The way in which objects are manipulated is highly related to their intrinsic properties. Here we used representational similarity analyses to investigate which object properties (i.e., elongation, size, or shape) are most strongly encoded and whether this information is coded together with motor aspects of the grasping task (i.e., degree of precision or number of digits). Since the similarity matrices revealed differences among conditions during passive viewing and grasping in most of the regions, we will discuss below results separately during the two task modalities. Since similar regions were identified using the searchlight and the ROI analyses, we will focus the discussion on the results of the latter.

Visuomotor gradient during grasping

Based on their neural similarity, regions grouped in the following clusters during grasping: one cluster included M1, S1, and the premotor cortex (PMd/FEF and PMv) and one cluster comprised visual areas (V1, LOC, pMTG) and parietal areas (aSPOC, pSPOC, pIPS, mIPS, aIPS). To focus on the neural similarity within regions of the same cluster, we discuss the two clusters separately.

Motor and premotor areas code number of digits

Motor and premotor areas showed a clear preference for motor coding in the grasping task, except for the PMd/FEF, where num-

ber of digits, object elongation, and size were represented to a similar extent. The homologous macaque area, F2, is mostly known to be involved in the transport phase of the reaching movement (Caminiti et al., 1991), but it has also been shown to be involved in grasping (Raos et al., 2004). Similarly, the human PMd showed directional tuning as well as grip selectivity (Fabbri et al., 2014) and adaptation for grasp-relevant object dimensions (Monaco et al., 2014). Our results support the possible involvement of the PMd in the continuous updating of the configuration and orientation of the hand as it approaches the object to grasp (Raos et al., 2004). S1, M1, and the PMv represented the number of digits. In S1, this sensitivity probably reflects the somatosensory stimulation elicited by different hand configurations. The finding that the PMv represented the motor task but none of the visual dimensions appears in contrast with object selectivity reported in the homologous monkey area, F5 (Murata et al., 1997). However, the covariation of object and grip type in that study did not allow to distinguish between area F5's role in coding grip and its role in coding object. Indeed, studies disambiguating grip selectivity and object type showed that the majority of neurons in this area were tuned to grip type and not object orientation (Fluet et al., 2010). Similarly, multiunit recordings in monkey area F5 found that it coded for grip type more than object orientation (Townsend et al., 2011). The strong motor representation in the PMv and the existence of direct connections between this area and M1 (Matelli et al., 1986) suggest a fundamental role of the PMv in the generation of grasping movements. The representation of the number of digits in M1 probably indicates that the index and the thumb are often used together to grasp, as well as all digits of the hand, in line with a recent study showing that M1 represents the use of fingers more than the individual fingers or their muscles (Ejaz et al., 2015). Our findings are in line with a study showing higher activation in M1, S1 and the PMv during Precision grip with three and five digits compared with two digits (Cavina-Pratesi et al., 2007*b*).

Visual and parietal areas encode object elongation and number of digits

V1, the aSPOC, and the pSPOC showed a clear preference for object elongation above the other visual and motor aspects. This sensitivity in V1 probably reflects different visual stimulation due to different object elongations, as this area is known to be selective for line orientation (Hubel and Wiesel, 1959). Even though the Elongation Model shows a winner-take-all pattern above all the other models in V1, it is at first surprising that motor codes were found in this area. In particular, Precision was coded more strongly than the Number of Digits. However, since grasping was executed in light, the two precision grasps led to a very similar visual configuration with the hand high on top of the object in both cases. It is likely that this visual similarity, more than a motor code per se, explains the significance of the Precision Model in V1. We found exclusive sensitivity to object elongation also in the pSPOC and the aSPOC, which have been proposed to be the human homolog of monkey area V6 and V6A, respectively (Cavina-Pratesi et al., 2010). Even though area V6A is part of the dorsolateral stream dedicated to the reaching component of the reach-to-grasp action (Jeannerod, 1999), this area showed selectivity for object type both during passive viewing and grasping (Fattori et al., 2010, 2012). Consistent with these results, the aSPOC showed adaptation for action-relevant features of the object instead of the visual size during grasping (Monaco et al., 2014). In light of these findings, the sensitivity for object elongation we found in the aSPOC might indicate that this specific

dimension is important for the visual extraction of object affordances and selection of wrist orientation, which is particularly important for elongated objects. We measured similar sensitivity in the pSPOC that correspond to monkey area V6, providing visual information to the aSPOC through direct connections (Galletti et al., 2001).

Differently from this winner-take-all pattern, the remaining regions represented both motor and visual aspects of the grasping task. In particular, in the pMTG and LOC, both visual and motor dimensions were represented to a similar extent, suggesting that object properties in these ventral regions are represented in relation to the grasping action. Previous studies also showed an involvement of ventral-stream areas in visually guided grasping (Culham et al., 2003; Cavina-Pratesi et al., 2010) and during hand-movement tasks based exclusively on somatosensory input (Fiehler et al., 2008). Recent studies also showed ventral-stream regions representing action-related object features, like the graspable dimension (Monaco et al., 2014) and object weight (Gallivan et al., 2014). Overall, these results are consistent with possible exchange of information between ventral and dorsal streams that controls skilled grasps (van Polanen and Davare, 2015) through the anatomical connections, such as those between the pMTG and aIPS (Borra et al., 2008). The aIPS in our study similarly coded number of digits, object elongation, and shape, suggesting that sensitivity for the hand configuration in this region is highly related to the object properties. These results apparently contrast with a preference for Precision 2 compared with Coarse 5 in previous univariate studies (Begliomini et al., 2007). However, multivariate analysis is more sensitive to detect representations distributed across voxels (Kriegeskorte et al., 2006). Indeed, a multivariate study showed the possibility of decoding different grip types in the aIPS (Di Bono et al., 2015). Our results suggest that the distinction between grip types arises from the number of digits more than the degree of precision required. The representation of object elongation in the pIPS and mIPS, along with the number of digits, agrees with neural selectivity for flat or elongated objects reported in the monkey cIPS (for review, see Sakata et al., 1997), a putative homolog of the human pIPS.

Changes in object representation during passive viewing and grasping

The representation of object properties differed between passive viewing and grasping. As expected by the specialized role of the ventral stream in “vision for perception” (Goodale and Milner, 1992), we found representations of object elongation and shape in V1 (and other low-level and mid-level visual areas) and the pMTG, while object shape alone was coded in the LOC. Object properties were also processed in parietal areas: object shape was represented in the aSPOC, while object elongation was represented in the mIPS and M1, as well as in the pSPOC, together with shape. Selectivity for object properties during passive viewing has been also reported in macaque parietal regions cIPS (Sakata et al., 1997) and area V6A (Breveglieri et al., 2015), probably reflecting the extraction of object-appropriate grips. Interestingly, the aIPS coded object shape and elongation during grasping but not passive viewing. Coding for object shape both in the aIPS and S1 might suggest sensitivity in the aIPS to different somatosensory information during manipulation of square and round objects, in line with the involvement of this area in integrating visual, sensorimotor, and motor information (Lewis and Van Essen, 2000). Object elongation was also coded in M1 during passive viewing, potentially reflecting unexecuted associations with the appropriate grasps.

In conclusion, by orthogonally manipulating object properties and grip types, we identified the most strongly represented visual and motor dimensions during grasping. In particular, we found that object elongation was a relevant property of the object to grasp, and that it was coded more strongly than size and shape in various human brain regions. The preference for object elongation during grasping but not during passive viewing might reflect the importance of this dimension in determining the optimal grip type and wrist orientation. The representation of object elongation, together the number of digits in both ventral-stream and dorsal-stream regions, suggests that communication between the two streams about these specific visual and motor dimensions is relevant to the execution of efficient grasping actions.

References

- Amiez C, Kostopoulos P, Champod AS, Petrides M (2006) Local morphology predicts functional organization of the dorsal premotor region in the human brain. *J Neurosci* 26:2724–2731. [CrossRef Medline](#)
- Baumann MA, Fluet MC, Scherberger H (2009) Context-specific grasp movement representation in the macaque anterior intraparietal area. *J Neurosci* 29:6436–6448. [CrossRef Medline](#)
- Begliomini C, Caria A, Grodd W, Castiello U (2007) Comparing natural and constrained movements: new insights into the visuomotor control of grasping. *PLoS One* 2:e1108. [CrossRef Medline](#)
- Borra E, Belmalih A, Calzavara R, Gerbella M, Murata A, Rozzi S, Luppino G (2008) Cortical connections of the macaque anterior intraparietal (AIP) area. *Cereb Cortex* 18:1094–1111. [CrossRef Medline](#)
- Bracci S, Op de Beeck H (2016) Dissociations and associations between shape and category representations in the two visual pathways. *J Neurosci* 36:432–44. [CrossRef Medline](#)
- Bracci S, Cavina-Pratesi C, Ietswaart M, Caramazza A, Peelen MV (2012) Closely overlapping responses to tools and hands in left lateral occipito-temporal cortex. *J Neurophysiol* 107:1443–1456. [CrossRef Medline](#)
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Breviglieri R, Galletti C, Bosco A, Gamberini M, Fattori P (2015) Object affordance modulates visual responses in the macaque medial posterior parietal cortex. *J Cogn Neurosci* 27:1447–1455. [CrossRef Medline](#)
- Caminiti R, Johnson PB, Galli C, Ferraina S, Burnod Y (1991) Making arm movements within different parts of space: the premotor and motor cortical representation of a coordinate system for reaching to visual targets. *J Neurosci* 11:1182–1197. [Medline](#)
- Castiello U (2005) The neuroscience of grasping. *Nat Rev Neurosci* 6:726–736. [CrossRef Medline](#)
- Cavina-Pratesi C, Goodale MA, Culham JC (2007a) fMRI reveals a dissociation between grasping and perceiving the size of real 3D objects. *PLoS One* 2:e424. [CrossRef Medline](#)
- Cavina-Pratesi C, Connolly JD, Monaco S, McAdam TD, Milner D, Schenk T, Culham J (2007b) Which aspects of hand-preshaping does human AIP compute during visually guided actions? Evidence from event-related fMRI. Paper presented at the Annual Meeting of the Society for Neuroscience, San Diego, CA, November.
- Cavina-Pratesi C, Monaco S, Fattori P, Galletti C, McAdam TD, Quinlan DJ, Goodale MA, Culham JC (2010) Functional magnetic resonance imaging reveals the neural substrates of arm transport and grip formation in reach-to-grasp actions in humans. *J Neurosci* 30:10306–10323. [CrossRef Medline](#)
- Culham JC, Danckert SL, DeSouza JF, Gati JS, Menon RS, Goodale MA (2003) Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp Brain Res* 153:180–189. [CrossRef Medline](#)
- DeLoache JS, Pierroutsakos SL, Uttal DH, Rosengren KS, Gottlieb A (1998) Grasping the nature of pictures. *Psychol Sci* 9:205–210. [CrossRef](#)
- Di Bono MG, Begliomini C, Castiello U, Zorzi M (2015) Probing the reaching-grasping network in humans through multivoxel pattern decoding. *Brain Behav* 5:e00412. [CrossRef Medline](#)
- Ejaz N, Hamada M, Diedrichsen J (2015) Hand use predicts the structure of representations in sensorimotor cortex. *Nat Neurosci* 18:1034–1040. [CrossRef Medline](#)

- Fabbri S, Strnad L, Caramazza A, Lingnau A (2014) Overlapping representations for grip type and reach direction. *Neuroimage* 94:138–146. [CrossRef Medline](#)
- Fang F, He S (2005) Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat Neurosci* 8:1380–1385. [CrossRef Medline](#)
- Fattori P, Raos V, Breveglieri R, Bosco A, Marzocchi N, Galletti C (2010) The dorsomedial pathway is not just for reaching: grasping neurons in the medial parieto-occipital cortex of the macaque monkey. *J Neurosci* 30:342–349. [CrossRef Medline](#)
- Fattori P, Breveglieri R, Raos V, Bosco A, Galletti C (2012) Vision for action in the macaque medial posterior parietal cortex. *J Neurosci* 32:3221–3234. [CrossRef Medline](#)
- Fiehler K, Burke M, Engel A, Bien S, Rösler F (2008) Kinesthetic working memory and action control within the dorsal stream. *Cereb Cortex* 18:243–253. [CrossRef Medline](#)
- Fluet MC, Baumann MA, Scherberger H (2010) Context-specific grasp movement representation in macaque ventral premotor cortex. *J Neurosci* 30:15175–15184. [CrossRef Medline](#)
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33:636–647. [CrossRef Medline](#)
- Frey SH, Vinton D, Norlund R, Grafton ST (2005) Cortical topography of human anterior intraparietal cortex active during visually guided grasping. *Brain Res Cogn Brain Res* 23:397–405. [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](#)
- Galletti C, Gamberini M, Kutz DF, Fattori P, Luppino G, Matelli M (2001) The cortical connections of area V6: an occipito-parietal network processing visual information. *Eur J Neurosci* 13:1572–1588. [CrossRef Medline](#)
- Gallivan JP, McLean DA, Valyear KF, Pettypiece CE, Culham JC (2011) Decoding action intentions from preparatory brain activity in human parieto-frontal networks. *J Neurosci* 31:9599–9610. [CrossRef Medline](#)
- Gallivan JP, Cant JS, Goodale MA, Flanagan JR (2014) Representation of object weight in human ventral visual cortex. *Curr Biol* 24:1866–1873. [CrossRef Medline](#)
- Ganel T, Goodale MA (2003) Visual control of action but not perception requires analytical processing of object shape. *Nature* 426:664–667. [CrossRef Medline](#)
- Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:20–25. [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](#)
- Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J Physiol* 148:574–591. [CrossRef Medline](#)
- Jeannerod M (1999) Visuomotor channels: their integration in goal-directed prehension. *Hum Mov Sci* 18:201–218. [CrossRef](#)
- Jeannerod M, Arbib MA, Rizzolatti G, Sakata H (1995) Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci* 18:314–320. [CrossRef Medline](#)
- Kriegeskorte N, Kievit RA (2013) Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn Sci* 17:401–412. [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, Bandettini P (2008) Representational similarity analysis—connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. [CrossRef Medline](#)
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12:535–540. [CrossRef Medline](#)
- Lewis JW, Van Essen DC (2000) Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J Comp Neurol* 428:112–137. [CrossRef Medline](#)
- Mahon BZ, Milleville SC, Negri GA, Rumati RI, Caramazza A, Martin A (2007) Action-related properties shape object representations in the ventral stream. *Neuron* 55:507–520. [CrossRef Medline](#)
- Matelli M, Camarda R, Glickstein M, Rizzolatti G (1986) Afferent and efferent projections of the inferior area 6 in the macaque monkey. *J Comp Neurol* 251:281–298. [CrossRef Medline](#)
- Monaco S, Chen Y, Medendorp WP, Crawford JD, Fiehler K, Henriques DY (2014) Functional magnetic resonance imaging adaptation reveals the cortical networks for processing grasp-relevant object properties. *Cereb Cortex* 24:1540–1554. [CrossRef Medline](#)
- Murata A, Fadiga L, Fogassi L, Gallese V, Raos V, Rizzolatti G (1997) Object representation in the ventral premotor cortex (area F5) of the monkey. *J Neurophysiol* 78:2226–2230. [Medline](#)
- Murata A, Gallese V, Luppino G, Kaseda M, Sakata H (2000) Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol* 83:2580–2601. [Medline](#)
- Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N (2014) A toolbox for representational similarity analysis. *PLoS Comput Biol* 10:e1003553. [CrossRef Medline](#)
- Raos V, Umiltà MA, Gallese V, Fogassi L (2004) Functional properties of grasping-related neurons in the dorsal premotor area F2 of the macaque monkey. *J Neurophysiol* 92:1990–2002. [CrossRef Medline](#)
- Sakata H, Taira M, Kusunoki M, Murata A, Tanaka Y (1997) The Trends Neurosci Lecture. The parietal association cortex in depth perception and visual control of hand action. *Trends Neurosci* 20:350–357. [CrossRef Medline](#)
- Sakata H, Taira M, Kusunoki M, Murata A, Tanaka Y, Tsutsui K (1998) Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. *Philos Trans R Soc Lond B Biol Sci* 353:1363–1373. [CrossRef Medline](#)
- Sakuraba S, Sakai S, Yamanaka M, Yokosawa K, Hirayama K (2012) Does the human dorsal stream really process a category for tools? *J Neurosci* 32:3949–3953. [CrossRef Medline](#)
- Schaffelhofer S, Agudelo-Toro A, Scherberger H (2015) Decoding a wide range of hand configurations from macaque motor, premotor, and parietal cortices. *J Neurosci* 35:1068–1081. [CrossRef Medline](#)
- Shikata E, McNamara A, Sprenger A, Hamzei F, Glauche V, Büchel C, Binkofski F (2008) Localization of human intraparietal areas AIP, CIP, and LIP using surface orientation and saccadic eye movement tasks. *Hum Brain Mapp* 29:411–421. [CrossRef Medline](#)
- Snow JC, Pettypiece CE, McAdam TD, McLean AD, Stroman PW, Goodale MA, Culham JC (2011) Bringing the real world into the fMRI scanner: repetition effects for pictures versus real objects. *Sci Rep* 1:130. [CrossRef Medline](#)
- Snow JC, Strother L, Humphreys GW (2014) Haptic shape processing in visual cortex. *J Cogn Neurosci* 26:1154–1167. [CrossRef Medline](#)
- Snow JC, Squires SD, Stubbs KM, Culham JC (2015) fMRI reveals different activation patterns for real objects vs. photographs of objects. Paper presented at the Society for Neuroscience, Chicago, IL, October.
- Taira M, Mine S, Georgopoulos AP, Murata A, Sakata H (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp Brain Res* 83:29–36. [Medline](#)
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. New York: Thieme.
- Tomassini V, Jbabdi S, Klein JC, Behrens TE, Pozzilli C, Matthews PM, Rushworth MF, Johansen-Berg H (2007) Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *J Neurosci* 27:10259–10269. [CrossRef Medline](#)
- Townsend BR, Subasi E, Scherberger H (2011) Grasp movement decoding from premotor and parietal cortex. *J Neurosci* 31:14386–14398. [CrossRef Medline](#)
- Van Essen DC (1997) A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature* 385:313–318. [CrossRef Medline](#)
- van Polanen V, Davare M (2015) Interactions between dorsal and ventral streams for controlling skilled grasp. *Neuropsychologia* 79:186–191. [CrossRef Medline](#)
- Young MP (1992) Objective analysis of the topological organization of the primate cortical visual system. *Nature* 358:152–155. [CrossRef Medline](#)
- Yousry TA, Schmid UD, Alkadhi H, Schmidt D, Peraud A, Buettner A, Winkler P (1997) Localization of the motor hand area to a knob on the precentral gyrus—a new landmark. *Brain* 120:141–157. [CrossRef Medline](#)