



Action Categories in Lateral Occipitotemporal Cortex Are Organized Along Sociality and Transitivity

Citation

Wurm, Moritz F., Alfonso Caramazza, and Angelika Lingnau. 2016. "Action Categories in Lateral Occipitotemporal Cortex Are Organized Along Sociality and Transitivity." *The Journal of Neuroscience* 37 (3): 562–75. <https://doi.org/10.1523/jneurosci.1717-16.2016>.

Published version

<https://doi.org/10.1523/JNEUROSCI.1717-16.2016>

Link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:41384850>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.

Please share how this access benefits you. [Submit a story](#)

Action Categories in Lateral Occipitotemporal Cortex Are Organized Along Sociality and Transitivity

Moritz F. Wurm,^{1,2} Alfonso Caramazza,^{1,2} and Angelika Lingnau^{2,3,4}

¹Cognitive Neuropsychology Laboratory, Harvard University, Cambridge, Massachusetts 02138, ²Center for Mind/Brain Sciences, University of Trento, 38068 Rovereto (TN), Italy, ³Department of Psychology, Royal Holloway University of London, TW20 0EX Egham, Surrey, United Kingdom, and

⁴Department of Cognitive Sciences, University of Trento, 38068 Rovereto (TN), Italy

How neural specificity for distinct conceptual knowledge categories arises is central for understanding the organization of semantic memory in the human brain. Although there is a large body of research on the neural processing of distinct object categories, the organization of action categories remains largely unknown. In particular, it is unknown whether different action categories follow a specific topographical organization on the cortical surface analogously to the category-specific organization of object knowledge. Here, we tested whether the neural representation of action knowledge is organized in terms of nonsocial versus social and object-unrelated versus object-related actions (sociality and transitivity, respectively, hereafter). We hypothesized a major distinction of sociality and transitivity along dorsal and ventral lateral occipitotemporal cortex (LOTc), respectively. Using fMRI-based multivoxel pattern analysis, we identified neural representations of action information associated with sociality and transitivity in bilateral LOTc. Representational similarity analysis revealed a dissociation between dorsal and ventral LOTc. We found that action representations in dorsal LOTc are segregated along features of sociality, whereas action representations in ventral LOTc are segregated along features of transitivity. In addition, representations of sociality and transitivity features were found more anteriorly in LOTc than representations of specific subtypes of actions, suggesting a posterior–anterior gradient from concrete to abstract action features. These findings elucidate how the neural representations of perceptually and conceptually diverse actions are organized in distinct subsystems in the LOTc.

Key words: action observation; categories; fMRI; MVPA; occipitotemporal cortex

Significance Statement

The lateral occipitotemporal cortex (LOTc) is critically involved in the recognition of objects and actions, but our knowledge about the underlying organizing principles is limited. Here, we discovered a dorsal–ventral distinction of actions in LOTc: dorsal LOTc represents actions based on sociality (how much an action is directed to another person) in proximity to person knowledge. In contrast, ventral LOTc represents actions based on transitivity (how much an action involves the interaction with inanimate objects) in proximity to tools/artifacts in ventral LOTc, suggesting a mutually dependent organization of actions and objects. In addition, we found a posterior-to-anterior organization of the LOTc for concrete and abstract representations, respectively. Our findings provide important insights about the organization of actions in LOTc.

Introduction

To be able to interact with our environment, we need to recognize objects and understand the actions of others. How the brain

achieves this task has been researched intensively in the last decades. Research demonstrated that distinct object categories are represented in a systematic topographical organization in occipitotemporal cortex (OTC) (Chao et al., 1999; Konkle and Caramazza, 2013). By contrast, the representation of action categories is less well understood (Pillon and d'Honinchtun, 2011). In particular, it is unclear whether actions are organized topographically along certain salient dimensions.

Two arguments support this assumption. First, according to the domain-specific hypothesis, distinct neural substrates became evolutionarily adapted to process knowledge categories selectively for which perceptual and conceptual distinctions lead to behavioral benefits (Caramazza and Shelton, 1998). Neuropsychological distinctions were identified among evolutionarily sa-

Received May 27, 2016; revised Nov. 7, 2016; accepted Nov. 11, 2016.

Author contributions: M.F.W., A.C., and A.L. designed research; M.F.W. performed research; M.F.W. analyzed data; M.F.W., A.C., and A.L. wrote the paper.

This work was supported by the German Research Foundation (DFG Research Grant WU 767/1-1), the Provincia Autonoma di Trento, and the Fondazione Cassa di Risparmio di Trento e Rovereto. We thank Paola Menapace for assistance in preparing the stimulus material and Gilles Vannuscorps for helpful comments on the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Moritz F. Wurm, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138. E-mail: mwurm@fas.harvard.edu.

DOI:10.1523/JNEUROSCI.1717-16.2016

Copyright © 2017 the authors 0270-6474/17/370562-14\$15.00/0

lient object categories such as animals, conspecifics, plant life, and tools (for review, see Caramazza and Mahon, 2003). In the action knowledge domain, a similar specialization might have occurred as certain behavioral “inventions” emerged and recognition of these behaviors became relevant for survival; for example, the distinction between social versus nonsocial and object-related (transitive) versus object-unrelated (intransitive) actions. Following this account, the neural processing of action knowledge along these dimensions (sociality and transitivity, hereafter) should be exposed to evolutionary pressure resulting in category-specific adaptation, and thus segregation, of the respective neural substrates. The second argument proposes that the neuroanatomical organization of action knowledge is determined by constraints from associated object categories. Action recognition comprises object recognition, specifically the recognition of the acting agent and other agents or inanimate objects that might be involved in the respective action. Strikingly, there are systematic links between certain action and object categories: Social actions (e.g., *teach*, *compete*, *sell*) are linked to knowledge about animate objects (e.g., conspecifics and interpersonal relations), whereas transitive actions (e.g., *cut*, *sew*, *peel*) are linked to knowledge about tools and other inanimate objects. The neural representations of actions and objects might therefore determine each other based on connectivity-based constraints. Because the most salient distinction of object knowledge is observed between animate and inanimate objects (Caramazza and Mahon, 2003; Martin, 2007; Kriegeskorte et al., 2008b), it is tempting to assume a similar prominent distinction in the action domain between sociality and transitivity.

How could the neural organization of object and action knowledge be related to each other? Animate objects activate dorsolateral OTC (DLOT), as well as lateral fusiform gyrus in ventral OTC (VOTC), whereas inanimate objects activate ventrolateral OTC (VLOT), as well as medial fusiform/parahippocampal cortex in VOTC (Chao et al., 1999; Downing et al., 2006; Konkle and Caramazza, 2013). Likewise, human motion preferentially activates DLOT, whereas tool motion preferentially activates VLOT (Beauchamp et al., 2002, 2003). Consistent with this mapping, processing of socially relevant cues draws on the superior temporal sulcus (STS; Allison et al., 2000; Carter and Huettel, 2013). Conversely, during action recognition, the lateral OTC (LOT) is predominantly activated (in addition to prefrontal and parietal areas that are not in the main focus of the present study; Van Overwalle and Baetens, 2009; Caspers et al., 2010). However, the precise organization of actions in LOT remains unclear and is a matter of current debate (Lingnau and Downing, 2015). Here, we hypothesize that social action knowledge is represented in the vicinity of animate and/or social-related information in DLOT, whereas transitive action knowledge is represented in the vicinity of inanimate object information (i.e., artifacts) in VLOT. Because the VOTC also reveals a distinction along animacy and is, albeit less often, found to be activated during action observation (Gobbini et al., 2007; Caspers et al., 2010; Shultz and McCarthy, 2012), it is possible that VOTC reveals a distinction of social (fusiform gyrus) and transitive actions (parahippocampal cortex) too. To test these predictions, we used fMRI-based multivoxel pattern analysis (MVPA) and representational similarity analysis (RSA) to investigate the neural organization of actions from four categories spanning a 2D semantic space along sociality and transitivity.

Materials and Methods

Participants. Twenty-eight healthy adults (8 females; mean age, 27 years; age range, 19–42 years) volunteered to participate in the experiment. All participants were right-handed with normal or corrected-to-normal vision and no history of neurological or psychiatric disease. Participants gave written informed consent before participation in the study. The experimental procedures were approved by the ethics committee for research involving human participants at the University of Trento, Italy.

Stimuli. The stimulus set consisted of 24 exemplars of eight actions (192 action videos in total). Actions were selected from four categories: (1) change of possession (transitive/social): *give*, *take*; (2) object manipulation (transitive/nonsocial): *open*, *close*; (3) communication (intransitive/social): *agree*, *disagree*; and (4) body/contact action (intransitive/nonsocial): *stroke*, *scratch*. The criteria for this selection were the following: only manual actions, actions that take place in the same context, and actions that are performed without physically interacting with, but in the presence of, another person. We thereby ensured that between-category analyses capture category-specific differences while eliminating feature differences that are not essential for an action category. Furthermore, we ensured that, within each category, actions are perceptually similar with regard to movement kinematics and complexity. We thereby guaranteed that within-category MVPA relied on conceptual but not perceptual differences between the two actions of a category. In addition, by using 24 different exemplars for each action (see Fig. 1B), we increased the perceptual variance of the stimuli to ensure that MVPA relied on abstract action representations that generalize across perceptual information (Wurm and Lingnau, 2015; Wurm et al., 2016). Variance was induced by using various stimulus factors; that is, two different contexts (kitchen, office), three perspectives (right, center, left; relative to the table orientation), two different actors (female, male), and six different objects that were present or involved in the actions (kitchen context: sugar cup, honey jar, coffee jar; office context: bottle, pen box, aluminum box). The actress/actor sat on either the left or the right side and used her/his right or left hand for the action. Stimulus factors were balanced for each action.

The concrete action instantiations were implemented as follows: *Give*: the actor moved an object from her/his peripersonal space into the peripersonal space of the passive person; *Take*: the reverse of *give*, i.e., the actor moved an object from the passive person's peripersonal space into her/his own peripersonal space; *Open*: the actor changed an object's state from closed into open; *Close*: the reverse of *open*, i.e., the actor changed an object's state from open into closed. (both actions required various different kinematics based on different lid/cap types, screw, push/pull, flip); *Agree*: the actor made a gesture in the direction of the passive person that signals agreement with the passive person (thumbs up, forming a ring with index finger and thumb); *Disagree*: the actor made a gesture in the direction of the passive person that signals disagreement with the passive person (thumbs down, waving with index finger; note that the heterogeneity of gestures ensured that MVPA could not rely on concrete hand postures, but only on the associated communicative meaning); *Stroke*: using the palm of the hand, the actor touched the other arm or hand lightly and repeatedly, as with brushing movements; and *Scratch*: using the fingertips, the actor scraped or rubbed the other arm or hand as if to relieve itching.

Catch trials consisted of six exemplars of each of the eight actions that deviated from the original action (e.g., tilting or lifting an object, making a meaningless gesture, incomplete actions, etc.; 48 catch trial videos in total). Action videos were filmed using a Canon 5D Mark II camera and edited in iMovie (Apple) and MATLAB (The MathWorks, RRID: SCR_006826). All 240 videos were identical in terms of action timing; that is, the videos started with hands on the table, followed by the action, and ended with hands moving to the same position of the table. Object states (open, closed) and positions (in front of the actress/actor or the passive person) were balanced in such a way that actions could not be predicted from the setting before the action started. Edited videos were grayscale and had a length of 2 s (30 frames per second) and a resolution of 400 × 225 pixels.

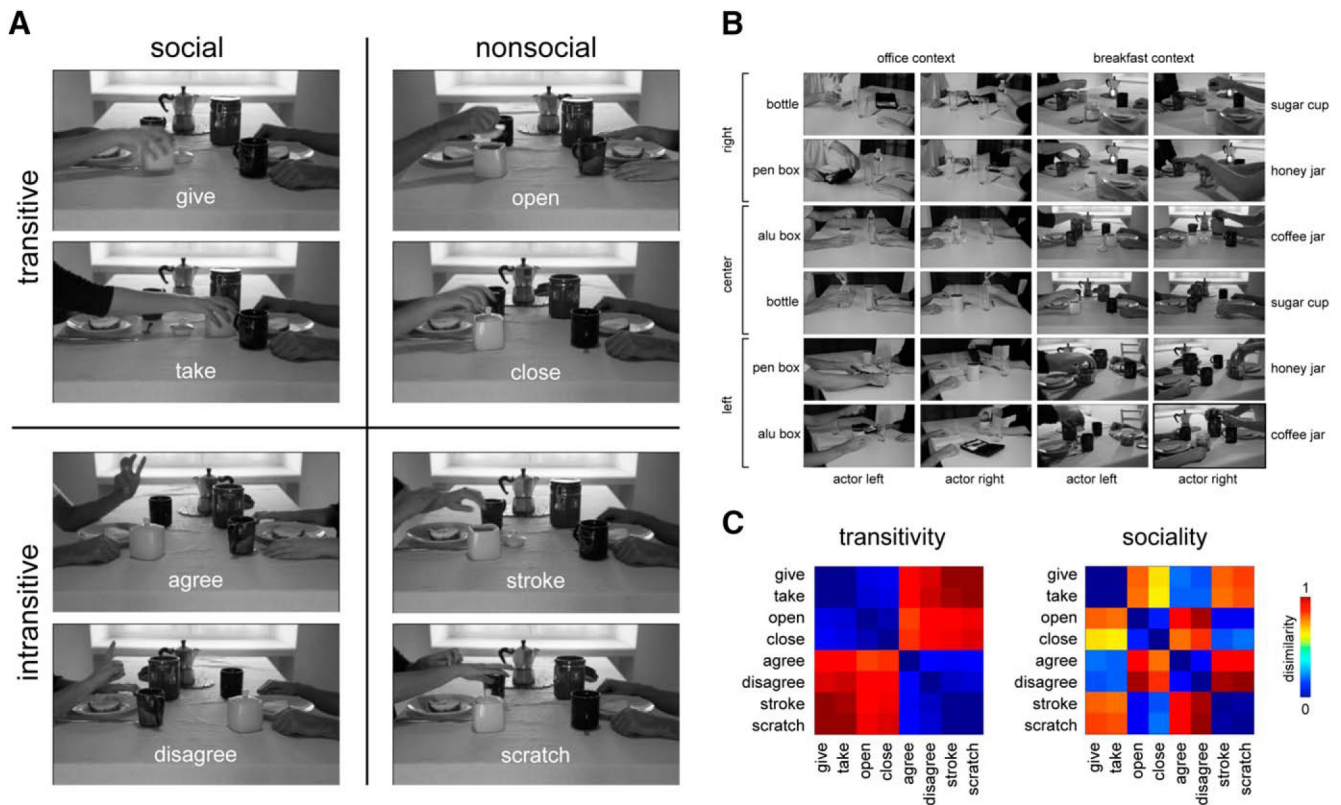


Figure 1. *A*, Experimental design with the factors transitivity and sociality using actions from four distinct categories (change of possession, object manipulation, communication, and body/contact actions). *B*, Stimulus variance (24 videos per condition). Actions were filmed from different perspectives, in different contexts, and involved different persons and objects to ensure that MVPA targets abstract action representations that generalize across various perceptual dimensions. *C*, RDMs for sociality and transitivity obtained from behavioral ratings (averaged across participants).

In the scanner, stimuli were back-projected onto a screen (60 Hz frame rate, 1024 × 768 pixels screen resolution) via a liquid crystal projector (OC EMP 7900, Epson) and viewed through a mirror mounted on the head coil (video presentation 6.9° × 3.9° visual angle). Stimulus presentation, response collection, and synchronization with the scanner were controlled with ASF (Schwarzbach, 2011) and the MATLAB Psychtoolbox-3 for Windows (Brainard, 1997).

Design of the fMRI experiment. Stimuli were presented in a mixed-event-related design. In each trial, videos (2 s) were followed by a 1 s fixation period. Eighteen trials were shown per block. Each of the nine conditions (eight action conditions plus one catch trial condition) was presented twice per block. Six blocks were presented per run, separated by 10 s fixation periods. Each run started with a 10 s fixation period and ended with a 16 s fixation period. In each run, the order of conditions was first-order counterbalanced (Aguirre, 2007). Each participant was scanned in a single session consisting of 8 functional scans and one anatomical scan. For each of the nine conditions there was a total of 2 (trials per block) × 6 (blocks per run) × 8 (runs per session) = 96 trials per condition. Each of the 24 exemplars per action condition was presented 4 times in the experiment.

Task. Participants were instructed to watch the movies attentively. They were asked to press a button with the right index finger on a response button box whenever an action was meaningless or performed incompletely or incorrectly (i.e., in catch trials). Participants could respond either during the movie or during the fixation phase after the movie. To ensure that participants followed the instructions correctly, they completed a practice run outside of the scanner. Participants were not informed about the exact purpose of the study and the organization of the actions into social/nonsocial and transitive/intransitive before the experiment.

After the fMRI session, participants judged the degree of sociality and transitivity of the actions seen in the experiment. To this end, 6-point

Likert scales (from 1 = not at all to 6 = very much) to the following questions were used: for transitivity, “How much does the action involve the interaction with a physical, inanimate object?”; for sociality, “How much is the action relevant for the nonacting person?” and “How much does the acting person consider possible consequences of the action for the nonacting person?” Ratings were used to ensure that the actions differed significantly along the two dimensions and were categorized as transitive/intransitive and social/nonsocial as intended. In addition, participants were asked to judge the similarity of the actions with regard to movement kinematics. We thereby ensured that sociality and transitivity are not confounded by covariance of movement differences between the actions. For each combination of the action conditions, participants judged on a 6-point Likert scale on how similar hand and arm movements of the respective actions were. Because different action instantiations were shown in the experiment, they were asked to focus on coarse-grained movements that were similar across the different instantiations. To test for covariance between sociality, transitivity, and movement similarity, we computed representational dissimilarity matrices (RDMs) by subtracting each rating value from each other (Euclidean distance) and used the vectorized triangle below the matrix diagonal for a correlation analysis for each participant. These RDM vectors were z-scored and correlated with each other to obtain one correlation coefficient (r) per correlation (sociality–transitivity, sociality–movement similarity, transitivity–movement similarity) and participant. We then used the r values in one-sample t tests to detect systematic correlations across participants. Averaged RDMs were used for RSA.

Data acquisition. Functional and structural data were collected using a 4 T Bruker MedSpec Biospin MR scanner and an 8-channel birdcage head coil. Functional images were acquired with a T2*-weighted gradient echoplanar imaging (EPI) sequence with fat suppression. Acquisition parameters were a repetition time of 2.2 s, an echo time of 33 ms, a flip angle of 75°, a field of view of 192 mm, a matrix size of 64 × 64, and a

voxel resolution of $3 \times 3 \times 3$ mm. We used 31 slices, acquired in ascending interleaved order, with a thickness of 3 mm and 15% gap (0.45 mm). Slices were tilted to run parallel to the superior temporal sulcus. In each functional run, 176 images were acquired. Before each run, we performed an additional scan to measure the point-spread function (PSF) of the acquired sequence to correct the distortion expected with high-field imaging (Zaitsev et al., 2004).

Structural T1-weighted images were acquired with an MPRAGE sequence (176 sagittal slices, TR = 2.7 s, inversion time = 1020 ms, FA = 7°, 256 × 224 mm FOV, 1 × 1 × 1 mm resolution).

Preprocessing. Data were analyzed using BrainVoyager QX 2.8 (Brain-Innovation, RRID:SCR_013057) in combination with the BVQXTools (RRID:SCR_009532) and NeuroElf (RRID:SCR_014147) toolboxes and custom software written in MATLAB (MathWorks).

Distortions in geometry and intensity in the EPIs were corrected on the basis of the PSF data acquired before each EPI scan (Zeng and Constable, 2002). The first four volumes were removed to avoid T1 saturation. The first volume of the first run was aligned to the high-resolution anatomy (six parameters). Data were 3D motion corrected (trilinear interpolation, with the first volume of the first run of each participant as reference), followed time correction and high-pass filtering (cutoff frequency of three cycles per run). Spatial smoothing was applied with a Gaussian kernel of 8 mm FWHM for univariate analysis and 3 mm FWHM for MVPA (Wurm and Lingnau, 2015). For group analysis, both anatomical and functional data were transformed into Talairach space using trilinear interpolation.

Univariate fMRI analysis. A group random-effects general linear model was computed using design matrices containing predictors of the 8 action conditions, catch trials, and of the 6 parameters resulting from 3D motion correction (x , y , z translation and rotation). Each predictor was convolved with a dual-gamma hemodynamic impulse response function (Friston et al., 1998). Each trial was modeled as an epoch lasting from video onset to offset (2 s). The resulting reference time courses were used to fit the signal time courses of each voxel. Statistical maps were thresholded using Threshold-Free Cluster Enhancement (TFCE; Smith and Nichols, 2009) as implemented in the CoSMoMVPA Toolbox (Oosterhof et al., 2016). A total of 10,000 Monte Carlo simulations and a corrected cluster threshold of $p = 0.05$ were used. Conjunctions were computed by outputting the minimum t value for each voxel of the input maps (Nichols et al., 2005). Maps were projected on a cortex-based aligned group surface for visualization.

MVPA. MVPA was performed using the CoSMoMVPA toolbox (Oosterhof et al., 2016). Design matrices contained 16 predictors reflecting the action conditions (8 actions × 2 exemplars), 2 catch trials predictors, and 6 predictors resulting from 3D motion correction. Beta weights of experimental conditions were estimated on the basis of six trials per condition and run resulting in two beta estimates per action condition and run. The 6 trials were selected from either the first half (blocks 1–3) or the second half (blocks 4–6) of each run. Because the six trials showed different instantiations of the same action (different contexts, perspectives, objects, actors, and hands), the MVPA targeted action representations that generalize across these factors. In total, this procedure resulted in 16 β maps (eight runs × two exemplars, hereinafter referred to as “patterns”) per action condition. Searchlight-based (Kriegeskorte et al., 2006) and ROI-based MVPA were performed in volume space using spherical ROIs with a radius of 12 mm. For searchlight analyses, individual accuracy maps were entered into a one-sample t test to identify voxels yielding classification significantly above chance. Statistical maps were corrected for multiple comparisons using TFCE (see “Univariate fMRI analysis” for details). For ease of comparison, the mean accuracy maps and the outlines of the corrected clusters were projected on the same cortex-based aligned group surface. Decoding analyses were performed using a linear discriminant analysis classifier.

Multiclass decoding. For multiclass searchlight MVPA, all eight actions were fed into the classification. In eight iterations, each action was discriminated from the remaining seven actions. The decoding accuracy at chance was therefore 12.5%. For within-category MVPA, only the two actions of the same category were decoded (e.g., open vs close for the transitive/nonsocial action category). The decoding accuracy at chance

was 50%. For both analyses, classification accuracies were computed using leave-one-out cross validation; that is, the classifier was trained using the data of 15 patterns and tested on its accuracy at classifying the unseen data from the remaining pattern. This procedure was performed in 16 iterations using all possible combinations of training and test patterns. The classification accuracies from the 16 iterations were averaged to give a mean accuracy score, which was assigned to the central voxel.

Across-category decoding. For the decoding of sociality and transitivity, the β values of the two actions were collapsed within each category. A cross-decoding scheme was used: To decode actions along transitivity, we trained the classifier to discriminate between transitive versus intransitive actions for the social dimension (give/take vs agree/disagree) and tested the classifier in the nonsocial dimension (open/close vs stroke/scratch). To decode actions along sociality, we trained the classifier to discriminate between social versus nonsocial actions for the transitive dimension (give/take vs open/close) and tested the classifier in the intransitive dimension (agree/disagree vs stroke/scratch). Both tests were also done vice versa (i.e., train on intransitive and test on the transitive dimension, train on nonsocial and test on the social dimension) and the resulting accuracies were averaged across the generalization directions. As described above, classification accuracies were computed using leave-one-out cross validation.

RSA. For RSA (Kriegeskorte et al., 2008a), we averaged the 16 β values of each action condition for each participant and voxel. For each searchlight/ROI sphere, we extracted the mean β values to obtain one multi-voxel pattern per action. For each pattern, we normalized the β values by subtracting the mean β value from each individual β value (demeaning). Next, we correlated the patterns with each other resulting in an 8×8 correlation matrix (the neural RDM) per sphere and participant. Then, neural RDMs were correlated with the RDMs for sociality and transitivity derived from the behavioral ratings. Resulting correlation coefficients were Fisher transformed and entered into one-sample t tests. Statistical maps were corrected for multiple comparisons using TFCE (see “Univariate fMRI analysis” for details).

Vector-of-ROI analysis. To analyze the topographical organization in LOTC and VOTC with respect to the different analyses (multiclass, across- and within-category decoding, RSA, univariate effects), we conducted a vector-of-ROI analysis (Konkle and Caramazza, 2013). To this end, we defined dorsal and ventral anchor points [posterior superior temporal sulcus (pSTS) and parahippocampal cortex, PHC] in each hemisphere based on the peak coordinates of the univariate conjunctions of sociality and transitivity (see Fig. 6). The anchor points were connected with a straight vector on the flattened cortical surface. This vector thus fully spanned LOTC and VOTC along the dorsal-ventral axis from pSTS (expected to be sensitive to person-related information) to PHC (expected to be sensitive to inanimate objects). Along this vector, we defined a series of partially overlapping spherical ROIs (12 mm radius, centers spaced 3 mm). In each ROI, we conducted all analyses as reported above using identical parameters as in the whole-brain analysis. For each analysis and hemisphere, responses were plotted as a function of position along the dorsal-ventral axis. Notably, as we focus on multivariate effects and use the univariate responses for comparison purposes only, the definition of the ROI vector is independent from the main analyses of interest.

For a second vector-of-ROI analysis along the posterior-anterior axis, the anchor points were based on anatomical landmarks: The posterior end was defined as the early visual cortex at the occipital pole; the anterior end was defined as the mid of the middle temporal gyrus (MTG). As is described above, the anchor points were connected with a straight vector on the flattened cortical surface. Along this vector, we defined a series of partially overlapping spherical ROIs (12 mm radius, centers spaced 3 mm apart). In each ROI, we conducted MVP decoding as reported above using identical parameters as in the whole-brain analysis.

Hierarchical cluster analysis. For additional visualization, we computed dendrograms of mean neural RDMs of DLOT and VLOT using hierarchical cluster analysis. DLOT and VLOT RDMs were extracted from the vector-of-ROI analysis. To this end, we first defined the center of action-sensitive LOTC as the peak of the multiclass decoding (see Fig. 7) and then defined DLOT and VLOT ROIs dorsally and ventrally of that peak; that

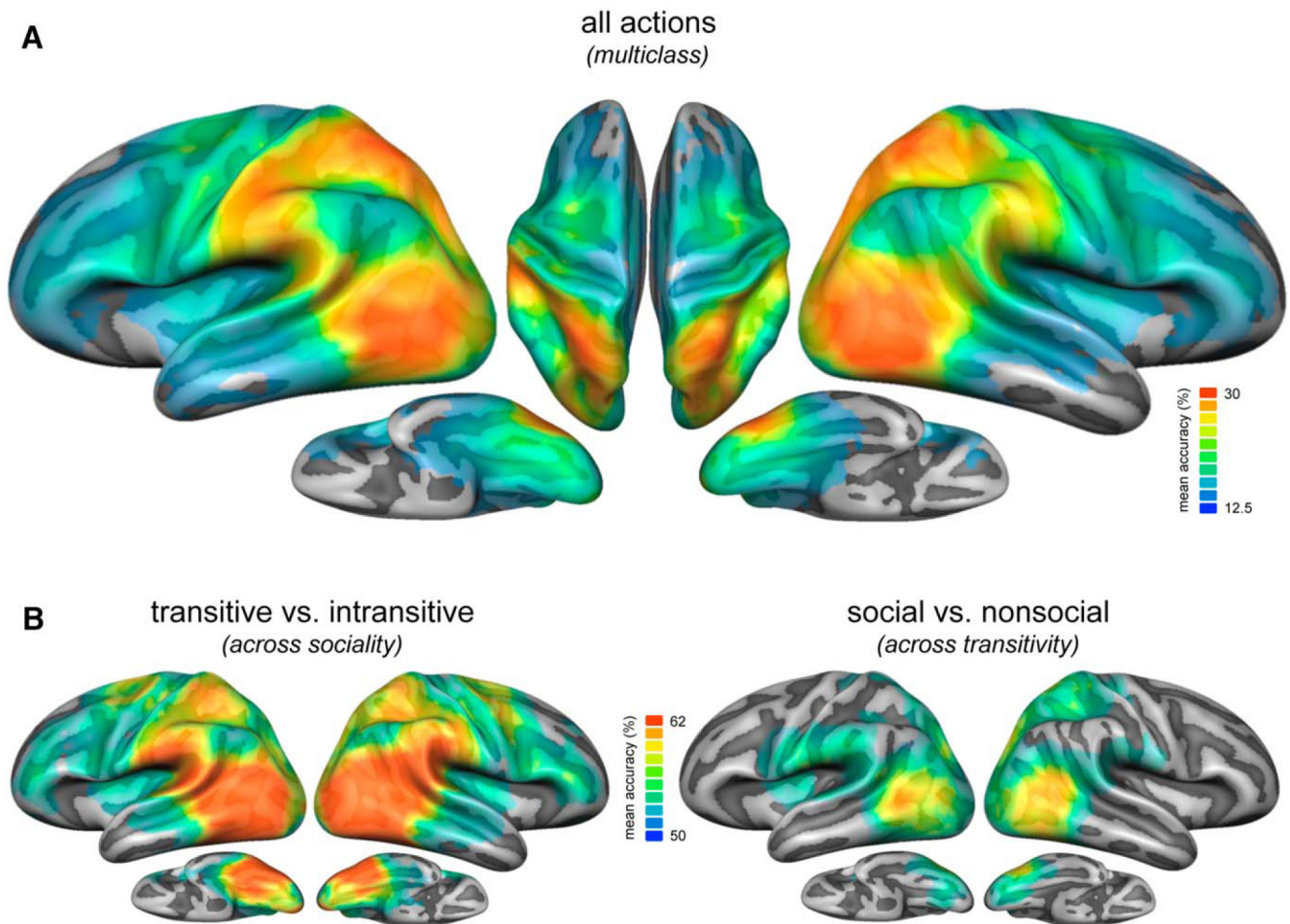


Figure 2. Mean accuracy maps of the searchlight-based multiclass decoding (each action against the remaining seven actions, chance = 12.5%; **A**) and the across-category decoding of sociality and transitivity (chance = 50%; **B**). Maps are thresholded using TFCE correction for multiple comparisons.

is, DLOTc was defined as eight adjacent ROIs dorsal of that peak and VLOTc was defined as eight adjacent ROIs ventral of that peak. For each hemisphere and ROI, RDMS were extracted and averaged across ROIs and participants. Hierarchical cluster analysis was performed using average distance.

Results

Behavioral results

All participants identified catch trials with sufficient accuracy, which ensured that participants paid attention to the action videos (mean error rates = $10.4 \pm 1.3\%$, SEM).

Behavioral ratings for sociality and transitivity revealed that actions were clearly categorized into transitive versus intransitive and social versus nonsocial, respectively (see Fig. 1C for the corresponding RDMS derived from the ratings). The two different ratings for sociality, which were sensitive for sociality with respect to the passive person or the actor, were strongly correlated with each other ($t_{(27)} = 11.3, p < 0.001$; mean $r = 0.99, p < 0.001$). We therefore collapsed the two ratings for subsequent analyses. Sociality and transitivity did not correlate significantly ($t_{(27)} = -0.082, p = 0.935$; mean $r = -0.066, p = 0.737$), which suggests that the two experimental dimensions were independent from each other. In addition, the two dimensions did not correlate with movement similarity (transitivity–movement similarity: $t_{(27)} = -0.144, p = 0.887$; mean $r = 0.088, p = 0.681$; sociality–movement similarity: $t_{(27)} = -0.935, p = 0.358$; mean $r = -0.291, p = 0.168$).

Brain regions sensitive to action discrimination

To get an overview of brain regions that are generally capable of discriminating actions of distinct categories, we performed a multiclass searchlight MVPA using all actions of the four categories. This analysis was sensitive to conceptual characteristics of the action categories (including sociality and transitivity) as well as to general movement types characteristic of the different categories (e.g., reaching/grasping, wrist rotation, hand gestures). Importantly, the high stimulus variance minimized the sensitivity to low-level perceptual differences between actions and maximized the sensitivity to action representations that generalize across features such as effector (right or left hand), perspective (view from left, right, or center positions; actor on the left or right side), and concrete movement (grasping/manipulating different objects; stroking/scratching different body parts; different gestures for agreement and disagreement, respectively).

The analysis revealed highly robust above chance decoding accuracies in lateral occipitotemporal and parietal regions that were strongest in left and right LOTc and middle intraparietal sulcus (IPS)/superior parietal lobe (SPL), respectively, as well as in the left posterior postcentral sulcus (PoCS)/anterior IPS (Fig. 2A). Decoding accuracies in frontal and medial temporal regions were substantially weaker than in the aforementioned regions (Table 1), which provides support to previous studies that found LOTc and PoCS/IPS, but less so premotor/prefrontal regions, to encode feature-general action representations (Oosterhof et al.,

Table 1. Clusters identified in multiclass action decoding and across-category decoding of sociality and transitivity

Region	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i>	<i>p</i>	Accuracy
All categories (multiclass decoding; chance = 12.5%)						
Left LOTC	−44	−64	3	13.88	8.28E−14	31.9
Right LOTC	44	−62	−5	15.65	4.66E−15	34.9
Left IPS/SPL	−21	−74	36	13.68	1.17E−13	27.4
Right IPS/SPL	23	−62	56	12.87	4.88E−13	29.1
Left PoCS/aIPS	−47	−27	37	13.67	1.19E−13	29.1
Right PoCS	58	−24	31	10.78	2.76E−11	25.7
Left PMv	−47	0	27	7.64	3.23E−08	20.8
Right PMv	35	−7	50	8.29	6.79E−09	22.3
Left PMd	−16	−7	58	9.19	8.61E−10	22.1
Left IFG	−50	10	18	8.34	5.98E−09	19.3
Sociality (across category decoding; chance = 50%)						
Left LOTC	−47	−58	4	9.32	6.24E−10	60.6
Right LOTC	48	−54	8	8.56	3.57E−09	60.5
Left TOS/IPS	−29	−69	20	7.59	3.67E−08	57.7
Right TOS/IPS	21	−81	33	8.11	1.03E−08	58.4
Left SMG	−58	−26	23	7.69	2.89E−08	56.8
Transitivity (across category decoding; chance = 50%)						
Left LOTC	−40	−61	−9	11.80	3.64E−12	67.8
Right LOTC	46	−55	−5	14.99	1.31E−14	69.9
Left SMG	−53	−36	25	11.36	8.78E−12	63.9
Right SMG	57	−24	26	11.79	3.71E−12	63.1
Left FG/PHC	−35	−46	−16	10.94	1.99E−11	65.2
Left PMd	−26	−20	61	9.01	1.26E−09	59.2
Right PMd	17	−9	62	8.86	1.78E−07	59.4
Left IFG	−39	32	19	6.16	1.34E−04	55.8

Shown are peak coordinates of corrected clusters in Talairach coordinates (*x*, *y*, *z*).

aIPS, anterior IPS; IFG, inferior frontal gyrus; IPMv, ventral premotor cortex; SMG, supramarginal gyrus; TOS, transverse occipital sulcus.

2012; Tucciarelli et al., 2015; Wurm and Lingnau, 2015; Wurm et al., 2016).

Brain regions sensitive to sociality and transitivity distinctions

Next, we investigated the functional organization of action representations with respect to sociality and transitivity features. In a first step, we searched for representations that are sensitive to sociality and transitivity independently of the concrete action subcategory. To this end, we performed an across-category decoding searchlight analysis. The general logic was the following: we trained a classifier to decode category A versus B and tested the same classifier using the categories C versus D (and vice versa). To decode sociality-specific features we trained a classifier to decode change of possession (trans/social) versus object manipulation (trans/nonsocial) and tested the classifier using communication (intrans/social) versus body/contact actions (intrans/nonsocial). Likewise, to decode transitivity-specific features, we trained a classifier to decode change of possession (trans/social) versus communication (intrans/social) and tested the classifier using object manipulation (trans/nonsocial) versus body/contact actions (intrans/nonsocial).

Both searchlight analyses revealed strong above chance accuracies in bilateral LOTC and, strikingly, far weaker effects in parietal regions (Fig. 2B). Decoding of sociality and transitivity features differed mostly with respect to overall decoding strength; that is, there were higher decoding accuracies for transitivity compared with sociality. This difference is not surprising because transitivity distinguishes actions based on salient perceptual features such as reaching

and grasping of objects, whereas sociality distinguishes actions based on more subtle, probably less perceptual, features. However, there were also anatomical differences. Although decoding of transitivity comprised regions in dorsal and ventral LOTC as well as in VOTC, decoding of sociality was mostly restricted to dorsal LOTC. Critically, in both hemispheres, decoding peaks of transitivity were in ventral LOTC, whereas peaks of sociality were in dorsal LOTC (see Fig. 5, Table 1).

In a second step, we characterized the representational organization of brain regions with respect to sociality and transitivity. To this end, we performed a searchlight-based RSA using RDMs obtained from behavioral ratings for sociality and transitivity (Fig. 3). The RSA for sociality revealed significant effects in bilateral pMTG, as well as in left postcentral gyrus. The RSA for transitivity revealed significant effects throughout lateral and VOTC (peaking in fusiform gyrus; FG/PHC), as well as in the posterior operculum, IPS, and PMd. Within LOTC, which has been suggested to be defined approximately by the boundaries middle portion of MTG (anterior), lateral occipital sulcus (posterior), STS (dorsal), and inferior temporal gyrus (ITG, ventral; Lingnau and Downing, 2015), the clusters found for sociality were located more dorsally than those for transitivity (Table 2). At a larger topographical scale, however, the dorsal–ventral gradient from transitivity to sociality was less strict because there were nearby regions fitting the transitivity model in regions other than ITG/FG/PHC (left posterior operculum/supramarginal gyrus and right posterior operculum/superior temporal gyrus), which were dorsal and anterior to the sociality clusters.

Brain regions representing category-specific subtypes of actions

The previous analysis focused on the abstract representation of sociality and transitivity features; that is, information that generalizes across category-specific actions. It is unclear, however, how these abstract dimensions relate neuroanatomically to more specific representations of action subtypes. To address this question, we decoded the actions for each category separately using a within-category searchlight MVPA (i.e., *give vs take*, *open vs close*, *agree vs disagree*, and *stroke vs scratch*). The critical difference between across-category and within-category MVPA is that the former relied on action-general differences between social versus nonsocial and transitive versus intransitive actions, respectively, whereas the within-category MVPA relied on action-specific differences between two actions of the same category. A notable feature of the within-category MVPA is that the decoded classes are perceptually similar so that the classifier exploits more subtle differences between actions: For example, videos of *give* and *take* contained highly similar reaching and grasping movements and differed only with respect to start and end location of the object relative to the actor [note that, due to the variance of actor position (left or right side of the table) and perspective, it is impossible that the decoding relied on absolute object positions]. Therefore, in the within-category MVPA, classification due to perceptual differences was minimized by keeping category-specific features such as reaching and grasping constant. In contrast, in the across-category MVPA, classification due to perceptual differences was minimized by generalizing across category-specific features. In addition, for both approaches, the high stimulus variance ensured that decoding relies on abstract representations that generalize across features such as effector and perspective.

In a first step, we performed searchlight analyses for each category separately. For each category, we obtained mean accu-

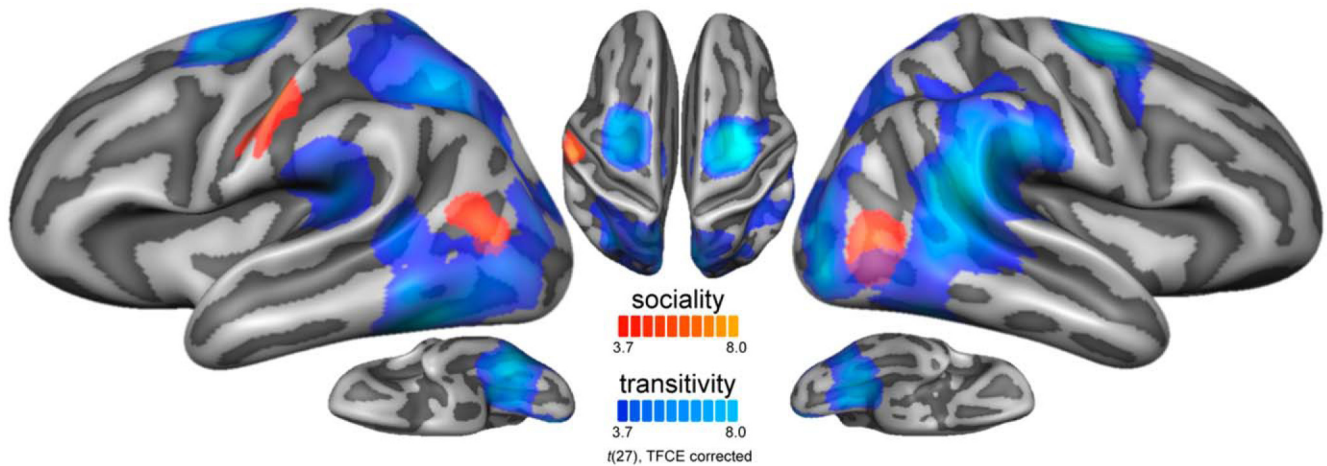


Figure 3. RSA. Statistical maps of the searchlight RSA for the transitivity model (blue) and the sociality model (red). Clusters are thresholded using TFCE correction for multiple comparisons.

Table 2. Clusters identified in searchlight RSA for sociality and transitivity

Region	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i>	<i>p</i>
Sociality RSA					
Left LOTC	−40	−66	9	4.75	6.02E−05
Right LOTC	43	−61	7	4.91	3.88E−05
Left PoCG	−49	−17	41	6.05	1.88E−06
Transitivity RSA					
Left VOTC/FG/PHC	−36	−45	−17	7.31	7.27E−08
Right VOTC/FG/PHC	24	−45	−6	6.65	3.87E−07
Right LOTC	38	−77	−3	7.32	7.14E−08
Left SMG/operculum	−41	−32	19	5.16	1.96E−05
Right SMG/operculum/STG	51	−40	15	7.86	1.87E−08
Left IPS/SPL	−29	−56	42	6.41	7.31E−07
Right SPL	7	−59	51	5.67	5.11E−06
Left PMd	−19	−10	51	6.94	1.83E−07
Right PMd	27	−10	49	8.43	4.86E−09
Left cuneus	−4	−85	2	7.65	3.19E−08
Right cuneus	4	−80	5	6.94	1.87E−07

Shown are peak coordinates of corrected clusters in Talairach coordinates (*x*, *y*, *z*).

PoCG, Postcentral gyrus; SMG, supramarginal gyrus.

acy and *t*-maps to reveal regions where decoding accuracy was consistently above chance (50%) across participants (Fig. 4A, Table 3). Decoding accuracies were generally highest in occipitotemporal and parietal regions; however, not all four searchlight analyses revealed statistically robust effects surviving TFCE correction. Decoding *open* versus *close* (object manipulation) revealed significant clusters in left LOTC and left postcentral sulcus. These clusters overlapped well with the clusters found in a previous study that decoded *open* versus *close* actions using different stimuli (Wurm and Lingnau, 2015). Decoding *agree* versus *disagree* (communication) revealed similar clusters in left and right LOTC and left postcentral sulcus ($p < 0.001$, uncorrected), but only the cluster in right LOTC survived TFCE correction. Decoding *give* versus *take* (change of possession) revealed a cluster in right LOTC ($p < 0.005$, uncorrected). Decoding *stroke* versus *scratch* (contact/body action) revealed a cluster in right precentral gyrus/sulcus ($p < 0.0025$, uncorrected). A comparison of the maps revealed no systematic segregation in LOTC along transitivity and/or sociality. This is perhaps not surprising because any higher-level information such as sociality and transitivity is constant between the decoded actions in the within-category MVPA and was thus canceled out. However, because the division into four separate searchlight analyses naturally reduced

the power for each analysis, we cannot rule out that also the representation of more concrete features of category-specific actions reflects distinctions along transitivity and/or sociality.

To investigate the general relationship between representations of the more abstract dimension sociality and transitivity with the more concrete representations of specific action subtypes regardless of the four categories, we collapsed the accuracy maps of the within-category MVPA for each participant and computed a *t* test across the averaged maps. We reasoned that this analysis should reveal areas containing representations of specific action subtypes regardless of the overarching action category. This analysis revealed significant clusters in left and right LOTC and left PoCS at the junction to the IPS (Fig. 4B). Clusters in left and right LOTC were more posterior to the clusters of the across-category decoding (Fig. 5). This finding points to a distinction between action-general and action-specific concept features along the anterior–posterior axis.

Although not the focus of the current study, the results obtained in parietal regions are worth mentioning. The clusters in PoCS partly overlapped with the anterior inferior parietal peak of the multiclass decoding. Interestingly, anterior IPL was found only to a weak extent in the across-category decoding. Consistent with previous findings (Oosterhof et al., 2010; Oosterhof et al., 2012; Leshinskaya and Caramazza, 2015; Wurm and Lingnau, 2015; Wurm et al., 2016), this pattern of results suggests that left anterior IPL represents action-specific information of a high degree of generality, but is less likely to represent higher-order dimensions such as sociality and transitivity. Anterior IPL thus reveals a functional profile that is different from the profile of LOTC and, notably, of posterior/superior parietal cortex (IPS/SPL). In IPS/SPL, effects were found only in the multiclass decoding and the RSA for transitivity but far less in the within- and across-category decoding. In other words, neural populations in IPS/SPL differentiate actions from one category from actions of other categories without generalizing across properties such as sociality and transitivity. At the same time, IPS/SPL did not differentiate actions of the same category when they were perceptually very similar; that is, have similar movement trajectories. Together, these findings suggest that IPS/SPL codes coarse-grained spatial action features specific for each of the categories. Consistent with studies on the role of IPS/SPL in action observation (Caspers et al., 2010; Binkofski and Buxbaum, 2013), it is likely that these features are related to body part motion in space

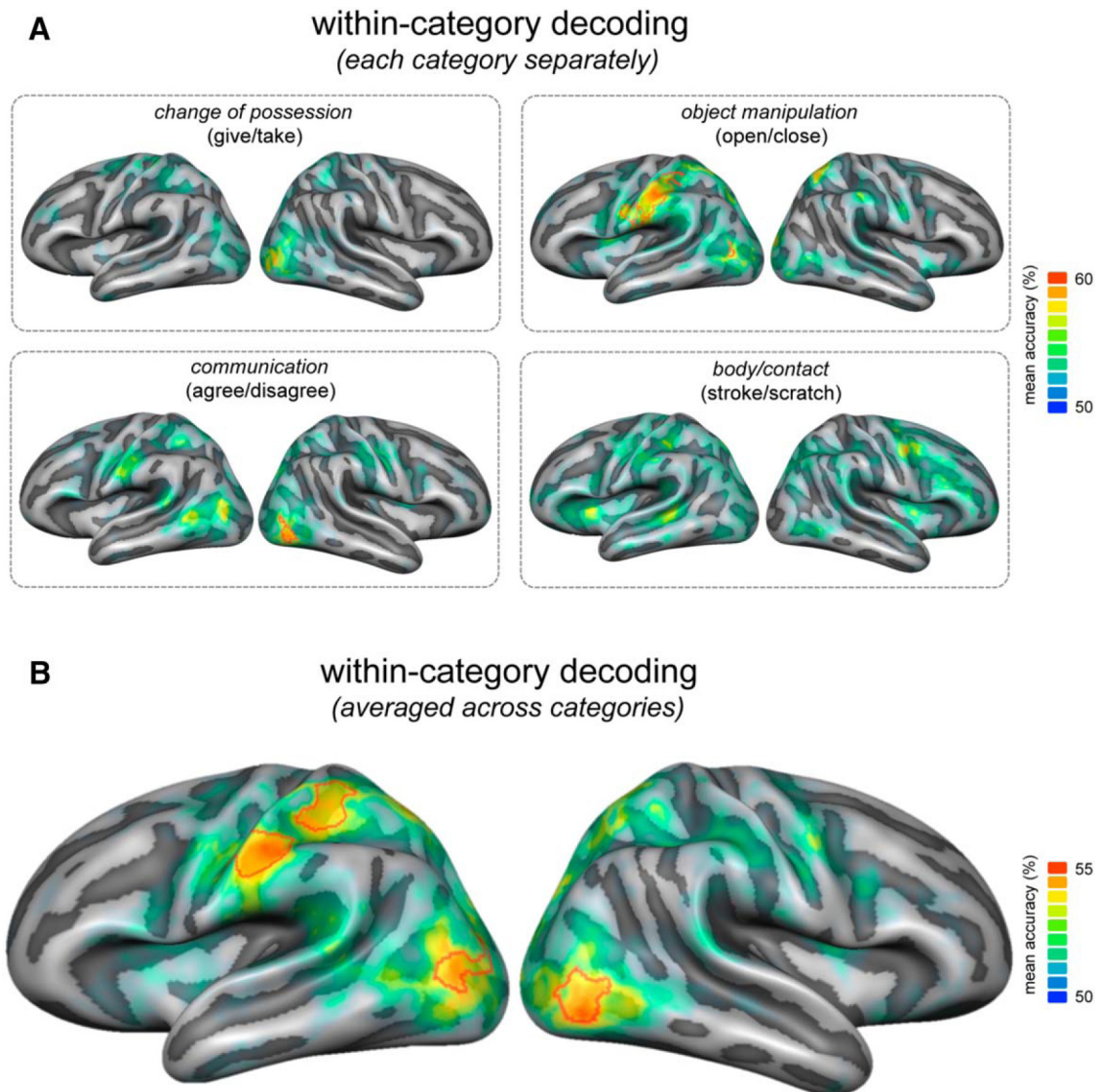


Figure 4. Mean accuracy maps for searchlight-based within-category decoding (chance = 50%) of each action category (**A**) and collapsed across categories (**B**). Clusters surviving TFCE correction for multiple comparisons are outlined in red.

that, in our study, was similar within category but different between categories.

Univariate (activation-based) effects of sociality and transitivity

Both across-category decoding and RSA analyses suggest distinct functional profiles in DLOTc and VLOTc regarding the action dimensions sociality and transitivity. Could this distinction be driven by increased activation of associated object information? For example, it is possible that the observation of social actions increased attention toward the nonacting person and thereby enhanced the processing of body and face information that could serve as socially relevant cues. Observation of social actions is also likely to induce mentalizing about the other person's feelings and reactions (Saxe and Kanwisher, 2003; Wurm et al., 2011). In contrast, observation of transitive actions is likely to direct attention toward the object involved in the action and thereby enhance the perceptual and semantic processing of that object. Following this logic, enhanced processing of person and inanimate object information should be reflected in enhanced activation in brain re-

gions representing person and inanimate object information, respectively. To identify regions showing such activation differences, we computed univariate contrast conjunctions for social versus nonsocial (*give/take* vs *open/close* and *agree/disagree* vs *stroke/scratch*) and transitive versus intransitive actions (*give/take* vs *agree/disagree* and *open/close* vs *stroke/scratch*), respectively (Fig. 6). The contrast conjunction social versus nonsocial revealed bilateral pSTS; that is, a region typically associated with the processing of socially relevant body and face information (Allison et al., 2000). Critically, in both hemispheres, the clusters in pSTS were dorsal to the DLOTc clusters identified in the sociality RSA. The contrast conjunction for transitive versus intransitive revealed bilateral FG/PHC, which can be associated with the processing of object information (Mahon et al., 2007), as well as the bilateral dorsal premotor cortex (PMd) and SPL; that is, regions recruited during the observation, planning, and execution of reaching and grasping movements (Binkofski and Buxbaum, 2013; Turella and Lingnau, 2014). These clusters overlapped with some of the clusters identified in the transitivity RSA, which suggests that, in these regions, multivariate effects might be affected

Table 3. Clusters identified in within-category decoding

Region	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i>	<i>p</i>	Accuracy
All categories (averaged)						
Left LOTC	−44	−73	5	5.15	2.01E−05	54.4
Right LOTC	42	−67	−5	5.68	4.84E−06	55
Left ventral PoCS	−54	−27	34	6.74	3.12E−07	54.9
Left dorsal PoCS	−32	−40	50	5.30	1.35E−05	54.1
Give versus take						
Right LOTC	40	−73	2	3.23	3.14E−03*	57.8
Right LOTC	31	−79	−3	3.39	2.09E−03*	56.9
Left PoCG	−31	−32	55	3.69	9.90E−04*	56.2
Open versus close						
Left LOTC	−44	−69	5	5.61	5.95E−06	57.2
Right LOTC	25	−83	11	5.04	2.73E−05*	57.7
Right MTG	48	−55	3	4.54	1.05E−04*	55.6
Left PoCS/aIPS	−51	−25	35	7.19	9.80E−07	59.6
Right SPL	27	−58	53	4.77	5.20E−05*	57.9
Agree versus disagree						
Right LOTC	42	−65	−4	6.44	6.69E−07	58.8
Left LOTC/MTG	−52	−52	4	4.47	1.15E−04*	56.6
Left LOTC	−42	−71	6	3.96	4.67E−04*	57.6
Left PoCS	−52	−24	35	4.01	4.06E−04*	55.9
Left IPS/SPL	−34	−46	45	3.79	7.29E−04*	55.4
Stroke versus scratch						
Right PMd	45	−4	43	3.79	7.37E−04*	58.1

Shown are peak coordinates of uncorrected (*) and corrected clusters in Talairach coordinates (*x, y, z*).

aIPS, anterior IPS; SMG, supramarginal gyrus.

Table 4. Clusters identified in univariate contrast conjunctions

Region	<i>x</i>	<i>y</i>	<i>z</i>	<i>t</i>	<i>p</i>
Social versus nonsocial					
Left pSTS	−46	−52	20	4.13	2.93E−04
Right pSTS	47	−63	23	3.91	5.38E−04
Transitive versus intransitive					
Left FG/PHC	−27	−44	−14	5.54	6.00E−06
Right FG/PHC	28	−45	−10	5.10	2.10E−05
Left PMd	−21	−8	51	5.07	2.30E−05
Right PMd	21	−7	52	6.22	1.00E−06
Left SPL	−30	−46	54	4.29	1.92E−04
Right SPL	28	−55	53	5.03	2.50E−05

Shown are peak coordinates of corrected clusters in Talairach coordinates (*x, y, z*).

by activation of inanimate object knowledge and kinematic representations. However, in LOTC, we did not find considerable activation differences overlapping with clusters identified in the transitivity RSA. Overall, the representational similarity seems to be rather independent from the univariate effects in LOTC (see also results of the vector-of-ROI analysis).

Vector-of-ROI analysis

To provide an integrated picture of the responses with respect to sociality and transitivity, we plotted multivariate effects, along with the univariate effects to each action category for reference, as a function of the position on a dorsal–ventral axis from the dorsal end of the LOTC (pSTS) to the ventral end of the VOTC (PHC). To this end, we defined anchor points based on the univariate contrast conjunctions for social versus nonsocial and transitive versus intransitive, respectively. These anchor points were chosen because we expected a putative segregation between transitive and social actions to be most eminent between regions sensitive to person-related information and inanimate objects, respectively. Between these anchor points, we defined a vector of adjacent ROIs. From each ROI, we extracted decoding accuracies (multiclass decoding, across- and within-category decoding),

RSA correlations, and univariate β estimates and plotted these responses as a function of the position on the dorsal–ventral axis. For univariate effects, we computed β estimates for each of the four action categories separately. For a better visualization of the relative differences between categories, we normalized β values (Konkle and Caramazza, 2013). For each ROI and category, we subtracted the mean of all four categories of that ROI. Results are shown in Figure 7. There were two major findings. First, the multiclass decoding (and to a less clear extent the within-category decoding) peaked in the LOTC at the level of MTG/ITG. This suggests that this region is generally most sensitive to action information. Second, the dorsal and ventral sides of this peak in LOTC showed preferences toward sociality and transitivity, respectively: across-category decoding and RSA revealed stronger effects of sociality on the dorsal compared with the ventral side of this peak. In contrast, effects of transitivity were stronger on the ventral compared with the dorsal side of this peak. These peaks were located between pSTS (dorsal end of the LOTC), the multiclass decoding peak in LOTC (middle of LOTC) and ITG (ventral end of LOTC; border to VOTC). Consistent with the univariate conjunction analysis (Fig. 6), pSTS and PHC showed univariate preferences for the two social (*give/take* and *agree/disagree*) and the two transitive (*give/take* and *open/close*) action categories, respectively. In addition, we observed a univariate preference for nonsocial action categories (*open/close* and *stroke/scratch*) in MTG/ITG, which could be due to increased processing of complex hand kinematics (Bracci et al., 2010; Orlov et al., 2014) that were specific for the two nonsocial action categories.

With regard to an additional segregation of sociality and transitivity in VOTC, the findings are less clear. The across-category decoding did not show systematic peak positions in FG and PHC that point to a distinction of sociality and transitivity. However, as expected, the RSA revealed a better fit of the sociality model in FG than in PHC, whereas the opposite effect was found for the transitivity model. It is questionable, though, whether this distinction reflects differences in representational organization of action knowledge because we did not observe a secondary peak of the multiclass decoding in VOTC, which should be the case if this region represented an additional hub of action processing.

Notably, along the whole dorsal–ventral axis, the across-category decoding of transitivity revealed higher accuracies than sociality, whereas the RSA showed higher correlations for sociality than for transitivity in DLOT (see also the respective searchlight analyses). This apparent discrepancy can be explained by the different methods underlying MVPA and RSA: using MVPA, the classifier might have picked up different (and possibly more subtle but highly reliable) information than the RSA (which is based on correlations of whole voxel patterns without biasing single, more reliable voxels).

The differential organization of action information along sociality and transitivity in LOTC was further illustrated by a hierarchical cluster analysis. In DLOT, social and nonsocial actions formed superordinate clusters; in VLOT, transitive and intransitive actions formed superordinate clusters (Fig. 8).

Finally, to investigate the gradient from action-specific to more general action features along the posterior–anterior axis, we conducted a second vector-of-ROI analysis. For investigation of this action-specificity gradient, only the within- and across-category decoding is informative. We therefore performed only these decoding analyses (Fig. 9). Consistent with the whole-brain analysis (Fig. 5), in both hemispheres, the peaks of the across-category decoding were located more anteriorly relative to the within-category decoding. Note however, that the within-category decoding revealed only subtle variations along the posterior–anterior axis; that is, there was no

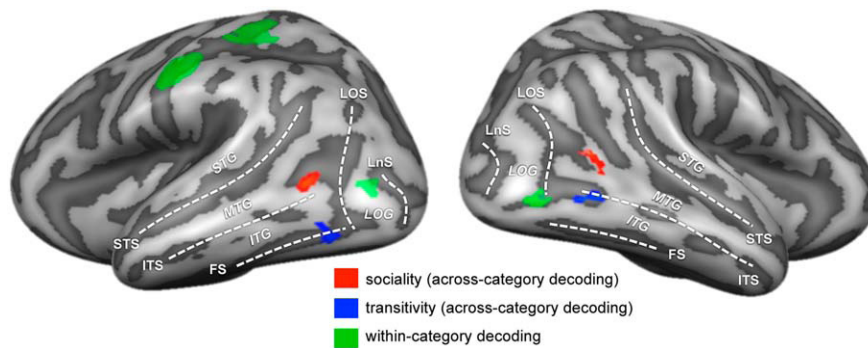


Figure 5. Peak clusters of the across-category decoding of sociality (red) and transitivity (blue) and of the within-category decoding (all categories collapsed; green). Peak clusters were created by identifying the peak vertex and adding nine adjacent vertices with the next highest t values in an iterative manner; that is, after adding the second vertex to the peak vertex, the third vertex adjacent to the two vertices with the highest t value was added, etc. FS, fusiform sulcus; ITS, inferior temporal sulcus; LNS, lunate sulcus; LOG, lateral occipital gyrus; LOS, lateral occipital sulcus.

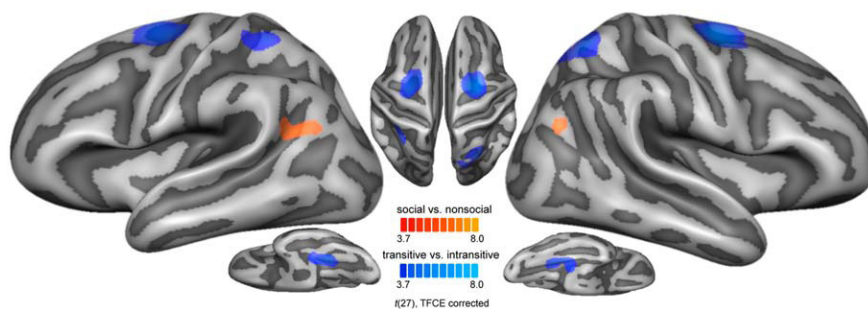


Figure 6. Conjunctions of the univariate contrasts for sociality (social/transitive vs nonsocial/transitive and social/intransitive vs nonsocial/intransitive; red) and transitivity (social/transitive vs social/intransitive and nonsocial/transitive vs nonsocial/intransitive; blue). Clusters are thresholded using TFCE correction for multiple comparisons.

clearly outstanding peak. This analysis therefore provided only moderate evidence for an action-specific to general gradient along the posterior axis.

Discussion

The present study investigated the neural organization of actions along the dimensions of sociality and transitivity. We report three major findings:

First, features associated with social versus nonsocial and transitive versus intransitive actions could be decoded in LOTC independently of the specific action category. For example, a classifier that was trained to distinguish between change of possession (social/transitive) and object manipulation (nonsocial/transitive) actions was able to distinguish between communicative (social/nontransitive) and body/contact (nonsocial/nontransitive) actions. This finding suggests that LOTC represents features of sociality and transitivity at a level that is independent of specific action subtypes.

Second, dorsal and ventral subregions of LOTC were organized preferentially along sociality and transitivity, respectively. The representational similarity of actions in DLOTC was better explained by the sociality model than by the transitivity model, whereas in VLOTC, the opposite pattern was found. This suggests that DLOTC represents social and nonsocial action features distinctly, whereas VLOTC represents transitive and intransitive action features distinctly.

Finally, information about specific actions of the same category could be decoded in regions of LOTC that were posterior to the regions coding sociality and transitivity. This finding suggests a second organization principle in LOTC;

that is, a gradient from posterior to anterior LOTC coding action-specific to more general category features independent of specific actions, respectively.

Dorsal and ventral LOTC/MTG differentiate social versus nonsocial and transitive versus intransitive action features, respectively

Using RSA, we demonstrated that DLOTC represents preferentially actions as predicted by the sociality model, whereas VLOTC represents preferentially actions as predicted by the transitivity model. In addition, in both hemispheres, the peak location of the social versus nonsocial action decoding was dorsal to the peak location of the transitive versus intransitive action decoding. Together, these findings show that action information along these dimensions is represented differentially in DLOTC and VLOTC. Overall, action decoding was highest at the level of MTG/ITG (Figs. 2, 7). In contrast, univariate effects of sociality and transitivity were found in pSTS and FG/PHC (Fig. 6), regions involved in the processing of person-related information (Allison et al., 2000) and inanimate objects (Chao et al., 1999; Mahon et al., 2007), respectively. Actions, even from distinct action categories such as those in our study, have structural similarities (typically involving the dynamic processing of motion and change, are typically intentional, etc.) and are therefore likely to be represented by neural substrates with similar computational properties (Kaas and Catania, 2002; Rosa and Tweedale, 2005). In other words, actions such as *open*, *give*, *agree*, and *scratch* are more similar to each other than to other, structurally different kinds of information such as persons and inanimate objects even if these kinds of information are important (albeit not constitutive) for action recognition. Following this reasoning, our finding that action information was encoded in proximity to, but not overlapping with, person-related and inanimate object knowledge is plausible. The subdivision within action-processing neural substrates along the dorsal–ventral axis (i.e., DLOTC) is more sensitive to sociality features, whereas VLOTC is more sensitive to transitivity features can be explained under the assumption that the neuro-anatomical organization of action knowledge is shaped by systematic connections between object and action representations: Socially relevant person information in dorsal areas such as the STS should be more strongly connected to social action representations in LOTC. In contrast, inanimate object information in ventral areas such as the ITG and FG should be more strongly connected to object-directed action representations in LOTC. The connections to person-related and inanimate object information thus might exert opposing constraints on the representations of social and transitive actions, which could drive the anatomical segregation in the observed way. This interpretation is supported by recent studies demonstrating enhanced functional connectivity specific for inanimate objects (artifacts and tools) between FG and a region in LOTC overlapping with the

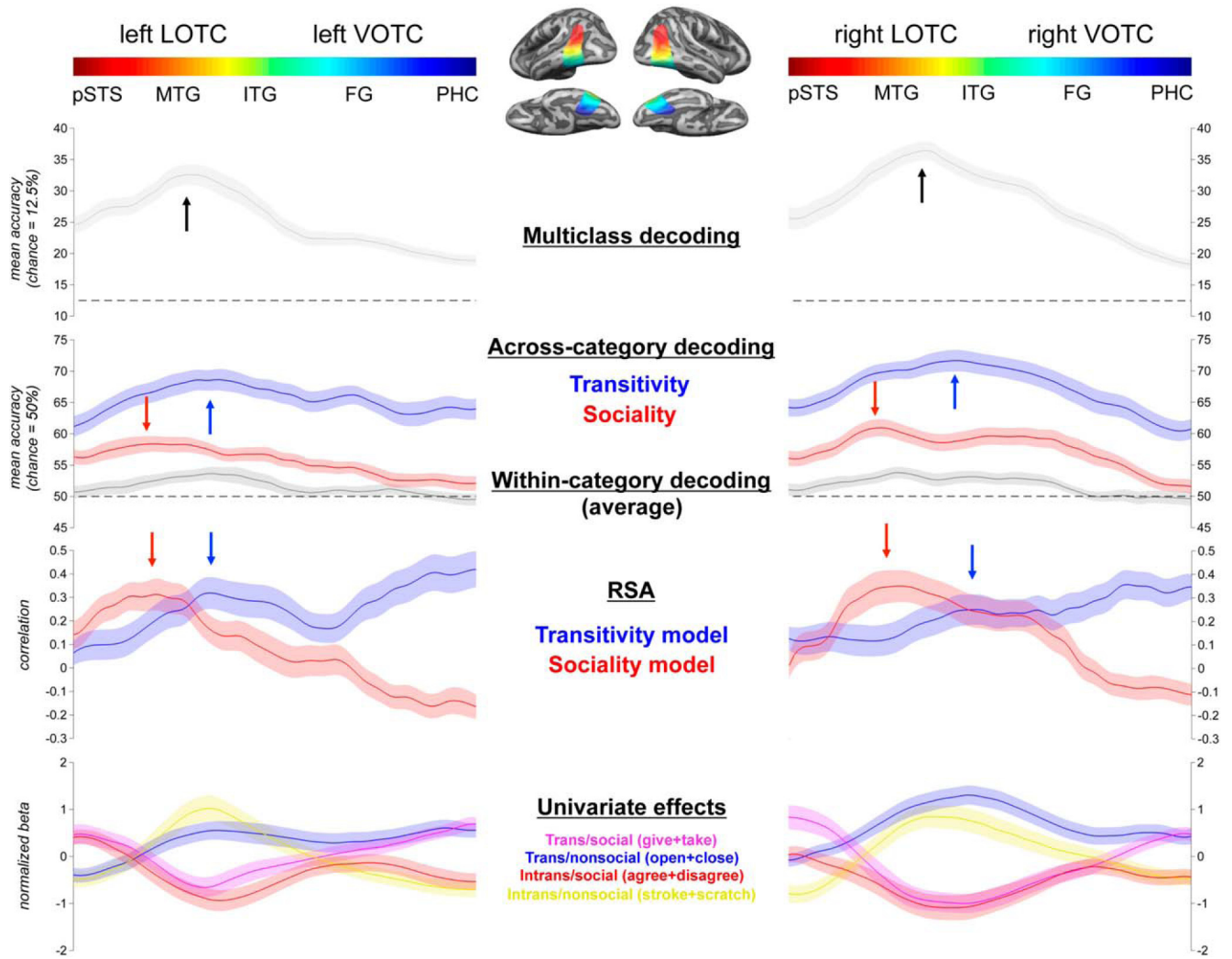


Figure 7. Vector-of-ROI analysis along the dorsal–ventral axis. To investigate the sociality–transitivity gradient, decoding accuracies (multiclass, across- category, and within- category decoding of transitivity and sociality), RSA correlations of neural RDMs with the transitivity and sociality models and univariate β estimates for each action category are plotted as a function of position on the dorsal–ventral axis from pSTS to PHC (see Materials and Methods for details). The color bar corresponds to the colors of the ROIs projected on the inflated cortex surface. Labels denote the approximate anatomical regions in LOTC and VOTC. Shaded areas around the curves represent 1 SEM across subjects. Black arrows indicate peaks of the multiclass decoding. Red and blue arrows indicate peaks in LOTC (located dorsally and ventrally of the multiclass peak) for sociality and transitivity, respectively, as revealed by the across-category decoding and the RSA.

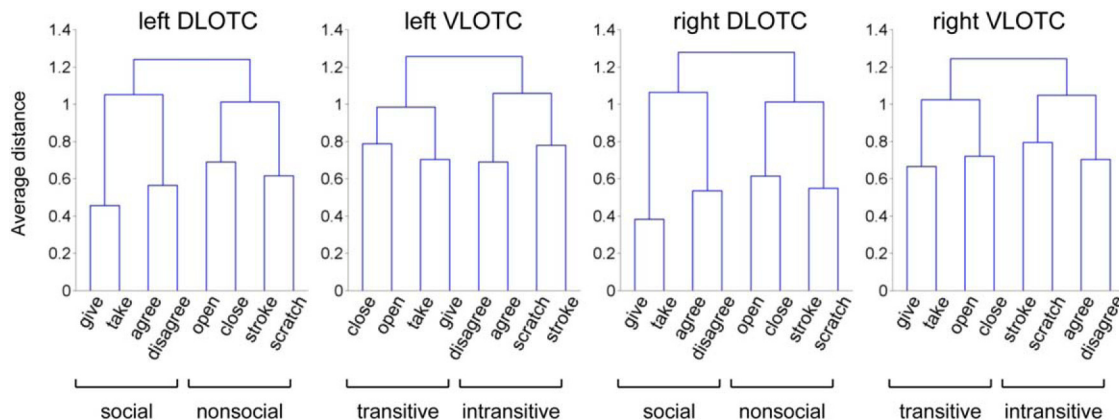


Figure 8. Dendrogram plots of the hierarchical cluster analysis (average distance) for dorsal and ventral LOTC. ROIs were defined by selecting eight adjacent ROIs of the vector-of-ROI analysis located dorsally (DLOTc) and ventrally (VLOTc) of the peak ROI of the multiclass decoding (Fig. 7).

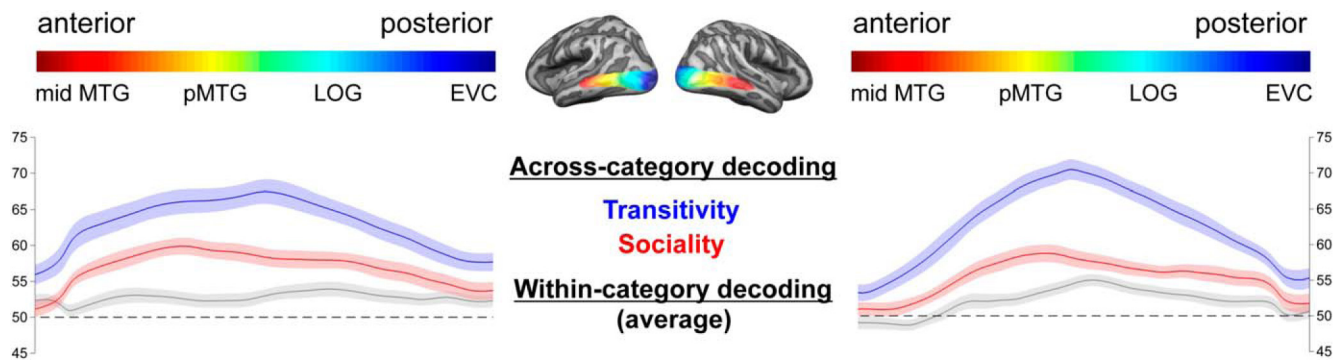


Figure 9. Vector-of-ROI analysis along the posterior–anterior axis. To investigate the gradient from action-specific to general action information, across- and within-category decoding accuracies are plotted as a function of position on the posterior–anterior axis from early visual cortex (EVC) to mid-MTG (see Materials and Methods for details). Color bars correspond to the colors of the ROIs projected on the inflated cortex surface. Labels denote the approximate anatomical regions from EVC to mid-MTG. Shaded areas around the curves represent 1 SEM across subjects.

region that we found to be sensitive for transitive versus intransitive action discrimination (Hutchison et al., 2014; Stevens et al., 2015). Likewise, effective connectivity between pSTS and LOTC is modulated by socially relevant cues such as facial expressions (Furl et al., 2015).

What remains unspecified is the kind of information that drives the observed distinctions in DLOT and VLOT as revealed by the RSA and the across-category decoding. Do the distinctions reflect semantic categorizations or are they driven by structural properties of the observed actions? Transitive actions can indeed be differentiated from intransitive actions based on intrinsic structural properties such as the reaching and grasping of objects. It is reasonable to assume that neural systems important for the recognition of reaching and grasping as well as hand–object interaction would be located in proximity to regions coding tools and other graspable objects (Bracci et al., 2012). The high structural similarities of actions within the transitive and intransitive categories are also reflected in the overall higher accuracies of the across-category decoding for transitivity. For social actions, perceptual commonalities are less evident. *Give* and *take* are perceptually different from *agree* and *disagree* gestures and, likewise, *open* and *close* are perceptually different from *stroke* and *scratch* actions. Consistent with this view, multivariate effects of sociality were generally more subtle than effects of transitivity. Furthermore, in both social and nonsocial actions, a second, passive person was present ruling out that social actions could be distinguished from nonsocial actions based on perceptual cues of the other person. Increased processing of the passive person for social actions is unlikely to drive the distinction because, in that case, one should also have observed univariate activation differences between social and nonsocial actions in LOTC, which was not the case. However, the social actions were directed to the other person and can thus be interpreted as interpersonal actions even if there was no observable reaction of the passive person. For the social actions, the acting and the passive person therefore defined a common social space, which was less the case for the nonsocial actions. The distinction between social and nonsocial could therefore be explained by more general underlying dimensions such as social space or the direction of an action toward another person or not. Another possibility is that general differences in the complexity of fine hand/finger movements, independently of the concrete movements themselves, drove the distinction between social versus nonsocial actions. Indeed, we found stronger univariate responses for the nonsocial versus social actions at the level of the pMTG (Fig. 7). Note, however, that the univariate response profile of the nonsocial

actions differed from the profile of the sociality RSA, which suggests that the two analyses picked up different kinds of information. Finally, it is possible that the across-category decoding relied on semantic representations of action primitives (Schank, 1973; Schank and Abelson, 1977) that were similar for the social actions and the transitive actions, respectively. In fact, the social actions used in our study (*give*, *take*, *agree*, *disagree*) involved a *transfer* of (physical or mental) objects. At the same time, the transitive actions (*give*, *take*, *open*, *close*) involved a change (of location or configuration) of objects. Action concepts that are composed of similar action primitives would therefore be close to each other in representational space, consistent with our findings. Future studies should investigate the extent to which such decompositional models (Jackendoff, 1972; Gruber, 1976; Pinker, 1989) can explain the neural organization of action knowledge.

Posterior to anterior LOTC is organized along a gradient from action-specific to general action information

A secondary finding of our study is that cluster peaks of the within-category decoding were located more posteriorly in LOTC than the peaks of the across-category decoding. Note however, that the range of accuracies of the within-category decoding was relatively shallow and the clusters of both decoding analyses showed substantial overlaps. The analysis therefore suggests only subtle, preferential differences of representational content along the posterior–anterior axis.

Compared with the across-category decoding, the within-category decoding relied on more subtle differences between actions of the same category (e.g., *give* vs *take* or *agree* vs *disagree*). These differences were either at a higher visual level (e.g., the position change of an object away vs toward the body of the acting person in the case of *give* vs *take*) or at the conceptual level (e.g., making different gestures for agreement vs disagreement in the case of *agree* vs *disagree*). The stimulus variance minimized the chance of decoding perceptual aspects of the actions such as perspective, agent, or concrete action instantiation. Because the actions were from the same category, it is not possible that decoding relied on more general features characteristic for an action category (e.g., transitivity for *open* vs *close* because both actions are transitive). In summary, the within-category decoding probably identified representations of specific action subtypes at a higher visual and/or conceptual level (Wurm and Lingnau, 2015).

In contrast, the across-category decoding was not suited to detect information specific for action subtypes because the classifier was trained and tested on actions of different categories (e.g., trained on change of possession vs object manipulation and tested on communication vs body/contact actions). As elaborated above, the

across-category decoding was most sensitive to action features that generalize across categories along the dimensions sociality and transitivity. The different peak locations of within- and across-category decoding suggest that abstract action-general features are represented more anteriorly than concrete action-specific features, which is consistent with recent proposals on the functional organization of LOTC from concrete to abstract and from visual to amodal action representations (see also Thompson-Schill, 2003; Martin, 2007; Watson and Chatterjee, 2011; Lingnau and Downing, 2015).

Conclusions

Our results suggest a topographic organization of LOTC along two major axes: a dorsal versus ventral distinction that segregates social versus object-related action information, respectively, and a posterior-to-anterior gradient from specific action subtypes to broader action categories that generalize across concrete action subtypes. This action topography gains its plausibility from the documented object topography, which distinguishes faces/bodies versus artifacts, and the connectivity of these object classes. Our results help to establish a clearer and theoretically motivated picture about the representational organization of LOTC.

References

- Aguirre GK (2007) Continuous carry-over designs for fMRI. *Neuroimage* 35:1480–1494. [CrossRef Medline](#)
- Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. *Trends Cogn Sci* 4:267–278. [CrossRef Medline](#)
- Beauchamp MS, Lee KE, Haxby JV, Martin A (2002) Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* 34:149–159. [CrossRef Medline](#)
- Beauchamp MS, Lee KE, Haxby JV, Martin A (2003) FMRI responses to video and point-light displays of moving humans and manipulable objects. *J Cogn Neurosci* 15:991–1001. [CrossRef Medline](#)
- Binkowski F, Buxbaum LJ (2013) Two action systems in the human brain. *Brain Lang* 127:222–229. [CrossRef Medline](#)
- Bracci S, Ietswaart M, Peelen MV, Cavina-Pratesi C (2010) Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *J Neurophysiol* 103:3389–3397. [CrossRef Medline](#)
- Bracci S, Cavina-Pratesi C, Ietswaart M, Caramazza A, Peelen MV (2012) Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *J Neurophysiol* 107:1443–1456. [CrossRef Medline](#)
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Caramazza A, Mahon BZ (2003) The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn Sci* 7:354–361. [CrossRef Medline](#)
- Caramazza A, Shelton JR (1998) Domain-specific knowledge systems in the brain the animate-inanimate distinction. *J Cogn Neurosci* 10:1–34. [CrossRef Medline](#)
- Carter RM, Huettel SA (2013) A nexus model of the temporal-parietal junction. *Trends Cogn Sci* 17:328–336. [CrossRef Medline](#)
- Caspers S, Zilles K, Laird AR, Eickhoff SB (2010) ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50:1148–1167. [CrossRef Medline](#)
- Chao LL, Haxby JV, Martin A (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2:913–919. [CrossRef Medline](#)
- Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N (2006) Domain specificity in visual cortex. *Cereb Cortex* 16:1453–1461. [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](#)
- Furl N, Henson RN, Friston KJ, Calder AJ (2015) Network interactions explain sensitivity to dynamic faces in the superior temporal sulcus. *Cereb Cortex* 25:2876–2882. [CrossRef Medline](#)
- Gobbini MI, Koralek AC, Bryan RE, Montgomery KJ, Haxby JV (2007) Two takes on the social brain: a comparison of theory of mind tasks. *J Cogn Neurosci* 19:1803–1814. [CrossRef Medline](#)
- Gruber J (1976) *Lexical structures in syntax and semantics*. New York: North-Holland.
- Hutchison RM, Culham JC, Everling S, Flanagan JR, Gallivan JP (2014) Distinct and distributed functional connectivity patterns across cortex reflect the domain-specific constraints of object, face, scene, body, and tool category-selective modules in the ventral visual pathway. *Neuroimage* 96:216–236. [CrossRef Medline](#)
- Jackendoff RS (1972) *Semantic interpretation in generative grammar*. Cambridge, MA: MIT.
- Kaas JH, Catania KC (2002) How do features of sensory representations develop? *BioEssays* 24:334–343. [CrossRef Medline](#)
- Konkle T, Caramazza A (2013) Tripartite organization of the ventral stream by animacy and object size. *J Neurosci* 33:10235–10242. [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 (2008a) Representational similarity analysis: connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. [CrossRef Medline](#)
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008b) Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60:1126–1141. [CrossRef Medline](#)
- Leshinskaya A, Caramazza A (2015) Abstract categories of functions in anterior parietal lobe. *Neuropsychologia* 76:27–40. [CrossRef Medline](#)
- Lingnau A, Downing PE (2015) The lateral occipitotemporal cortex in action. *Trends Cogn Sci* 19:268–277. [CrossRef Medline](#)
- Mahon BZ, Milleville SC, Negri GA, Rumiaty RI, Caramazza A, Martin A (2007) Action-related properties shape object representations in the ventral stream. *Neuron* 55:507–520. [CrossRef Medline](#)
- Martin A (2007) The representation of object concepts in the brain. *Annu Rev Psychol* 58:25–45. [CrossRef Medline](#)
- Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005) Valid conjunction inference with the minimum statistic. *Neuroimage* 25:653–660. [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](#)
- Oosterhof NN, Tipper SP, Downing PE (2012) Viewpoint (in)dependence of action representations: an MVPA study. *J Cogn Neurosci* 24:975–989. [CrossRef Medline](#)
- Oosterhof NN, Connolly AC, Haxby JV (2016) CoSMoMVPA: multimodal multivariate pattern analysis of neuroimaging data in MATLAB/GNU Octave. *Front Neuroinform* 10:27. [CrossRef Medline](#)
- Orlov T, Porat Y, Makin TR, Zohary E (2014) Hands in motion: an upper-limb-selective area in the occipitotemporal cortex shows sensitivity to viewed hand kinematics. *J Neurosci* 34:4882–4895. [CrossRef Medline](#)
- Pillon A, d'Honin P (2011) The organization of the conceptual system: the case of the “object versus action” dimension. *Cogn Neuropsychol* 27:587–613. [CrossRef Medline](#)
- Pinker S (1989) *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT.
- Rosa MG, Tweedale R (2005) Brain maps, great and small: lessons from comparative studies of primate visual cortical organization. *Philos Trans R Soc Lond B Biol Sci* 360:665–691. [CrossRef Medline](#)
- Saxe R, Kanwisher N (2003) People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage* 19:1835–1842. [CrossRef Medline](#)
- Shank RC (1973) *The fourteen primitive actions and their inferences*. Memo AIM-183, Stanford Artificial Intelligence Laboratory, Stanford University Stanford, CA.
- Shank RC, Abelson RP (1977) *Scripts, plans, goals, and understanding: an inquiry into human knowledge structures*. Hillsdale, NJ: Lawrence Erlbaum.
- Schwarzbach J (2011) A simple framework (ASF) for behavioral and neuroimaging experiments based on the psychophysics toolbox for MATLAB. *Behav Res Methods* 43:1194–1201. [CrossRef Medline](#)
- Shultz S, McCarthy G (2012) Goal-directed actions activate the face-sensitive posterior superior temporal sulcus and fusiform gyrus in the absence of human-like perceptual cues. *Cereb Cortex* 22:1098–1106. [CrossRef Medline](#)
- Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44:83–98. [CrossRef Medline](#)
- Stevens WD, Tessler MH, Peng CS, Martin A (2015) Functional connectiv-

- ity constrains the category-related organization of human ventral occipitotemporal cortex. *Hum Brain Mapp* 36:2187–2206. [CrossRef Medline](#)
- Thompson-Schill SL (2003) Neuroimaging studies of semantic memory: inferring “how” from “where”. *Neuropsychologia* 41:280–292. [CrossRef Medline](#)
- Tucciarelli R, Turella L, Oosterhof NN, Weisz N, Lingnau A (2015) MEG multivariate analysis reveals early abstract action representations in the lateral occipitotemporal cortex. *J Neurosci* 35:16034–16045. [CrossRef Medline](#)
- Turella L, Lingnau A (2014) Neural correlates of grasping. *Front Hum Neurosci* 8:686. [CrossRef Medline](#)
- Van Overwalle F, Baetens K (2009) Understanding others’ actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage* 48:564–584. [CrossRef Medline](#)
- Watson CE, Chatterjee A (2011) The functional neuroanatomy of actions. *Neurology* 76:1428–1434. [CrossRef Medline](#)
- Wurm MF, Lingnau A (2015) Decoding actions at different levels of abstraction. *J Neurosci* 35:7727–7735. [CrossRef Medline](#)
- Wurm MF, von Cramon DY, Schubotz RI (2011) Do we mind other minds when we mind other minds’ actions? A functional magnetic resonance imaging study. *Hum Brain Mapp* 32:2141–2150. [CrossRef Medline](#)
- Wurm MF, Ariani G, Greenlee MW, Lingnau A (2016) Decoding concrete and abstract action representations during explicit and implicit conceptual processing. *Cereb Cortex* 26:3390–3401. [CrossRef Medline](#)
- Zaitsev M, Hennig J, Speck O (2004) Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction. *Magn Reson Med* 52:1156–1166. [CrossRef Medline](#)
- Zeng H, Constable RT (2002) Image distortion correction in EPI: comparison of field mapping with point spread function mapping. *Magn Reson Med* 48:137–146. [CrossRef Medline](#)