The Journal of Neuroscience

http://jneurosci.msubmit.net

JN-RM-1717-16R2

Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity

Moritz Wurm, Harvard University
Alfonso Caramazza, Harvard University
Angelika Lingnau, Royal Holloway University of London

Commercial Interest:

Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity

Abbreviated title: Action categories in occipitotemporal cortex

Moritz. F. Wurm[1,2], Alfonso Caramazza[1,2], & Angelika Lingnau[2,3,4]

1: Cognitive Neuropsychology Laboratory, Harvard University, Cambridge, MA 02138, USA
2: Center for Mind/Brain Sciences, University of Trento, 38068 Rovereto (TN), Italy
3: Department of Psychology, Royal Holloway University of London, TW20 0EX Egham, Surrey, UK
4: Department of Cognitive Sciences, University of Trento, Corso Bettini, 31, 38068 Rovereto (TN), Italy

Corresponding author:
Moritz F. Wurm
Department of Psychology
Harvard University
33 Kirkland Street
Cambridge, MA 02138, USA
Email: mwurm@fas.harvard.edu

Number of pages: 42
Number of figures: 9
Number of Tables: 4

Number of words:
Abstract: 229
Introduction: 650
Discussion: 1640

Conflict of Interest: none

Acknowledgements: We would like to thank Gilles Vannuscorps for helpful comments on the manuscript. This research 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.

ABSTRACT

How neural specificity for distinct conceptual knowledge categories arises is central for understanding the organization of semantic memory in the human brain. While 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 if 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 non-social vs. social and object-unrelated vs. object-related actions (respectively, sociality and transitivity, 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 (MVPA), we identified neural representations of action information associated with sociality and transitivity in bilateral LOTC. Representational similarity analysis (RSA) 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.

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. By 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 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'Honincthun, 2011). In particular, it is unclear if actions are topographically organized along certain salient dimensions.

Two arguments support this assumption. First, according to the Domain-Specific Hypothesis, distinct neural substrates became evolutionary adapted to selectively process knowledge categories for which perceptual and conceptual distinctions lead to behavioral benefits (Caramazza and Shelton, 1998). Neuropsychological distinctions were identified among evolutionarily salient object categories like animals, conspecifics, plant life, and tools (see Caramazza and Mahon, 2003, for a review). 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, e.g. the distinction between social vs. nonsocial and object-related (transitive) vs. 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 as well as 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. Since 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 (DLOTC), as well as lateral fusiform gyrus in ventral OTC, whereas inanimate objects activate ventrolateral OTC (VLOTC), as well as medial fusiform/parahippocampal cortex in ventral OTC (Chao et al., 1999; Downing et al., 2006; Konkle and Caramazza, 2013). Likewise, human motion preferentially activates DLOTC whereas tool motion preferentially activates VLOTC (Beauchamp et al., 2002, 2003). In line with this mapping, processing of socially relevant cues draws on the superior temporal sulcus (STS) (Allison et al., 2000; Carter and Huettel, 2013). During action recognition, on the other hand, lateral OTC (LOTC) is predominantly activated (besides 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 LOTC 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 DLOTC, whereas transitive action knowledge is represented in the vicinity of inanimate object information (i.e., artifacts) in VLOTC. As ventral OTC 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 ventral OTC reveals a distinction of social (fusiform gyrus) and transitive actions (parahippocampal cortex) too. To test these predictions, we used fMRI-based multivoxel pattern analysis and representational similarity analysis to investigate the neural organization of actions from four categories spanning a two-dimensional semantic space along sociality and transitivity.

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 prior to 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: change of possession (transitive/social): *give*, *take*; object manipulation (transitive/nonsocial): *open*, *close*; communication (intransitive/social): *agree*, *disagree*; body/contact action (intransitive/nonsocial): *stroke*, *scratch*. The criteria for this selection were the following: (1) only manual actions, (2) actions that take place in the same context, and

(3) 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 (4), 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. Additionally, by using 24 different exemplars for each action (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 et al., 2015; Wurm and Lingnau, 2015). 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. *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 (MathWorks, RRID:SCR_006826). All 240 videos were identical in terms of action timing, i.e., 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 gray scale, had a length of 2 s (30 frames per second), and a resolution of 400 x 225 pixels.

In the scanner, stimuli were back-projected onto a screen (60 Hz frame rate, 1024 x 768 pixels screen resolution) via a liquid crystal projector (OC EMP 7900, Epson Nagano, Japan) and viewed through a mirror mounted on the head coil (video presentation 6.9° x 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. 18 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) x 6 (blocks per run) x 8 (runs per session) = 96 trials per condition. Each of the 24 exemplars per action condition was presented four times in the experiment.

Task. Participants were instructed to attentively watch the movies. 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 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: Transitivity: “How much does the action involve the interaction with a physical, inanimate object?” Sociality: “How much is the action relevant for the non-acting person?” and “How much does the acting person consider possible consequences of the action for the non-acting

person?" Ratings were used to ensure that the actions differed significantly along the two dimensions and were categorized as transitive/intransitive and social/non-social 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 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 dissimilarity matrices 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. The averaged dissimilarity matrices were also used as representational dissimilarity matrices (RDM) for representational similarity analysis (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 echo-planar 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 x 64, and a voxel resolution of 3 x 3 x 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-weigthed images were acquired with an MPRAGE sequence (176 sagittal slices, TR = 2.7 s, inversion time = 1020 ms, FA = 7°, 256 x 224 mm FOV, 1 x 1 x 1 mm resolution).

Preprocessing. Data were analyzed using BrainVoyager QX 2.8 (BrainInnovation, 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 echo-planar images were corrected on the basis of the PSF data acquired before each EPI scan (Zeng and Constable, 2002). The first 4 volumes were removed to avoid T1 saturation. The first volume of the first run was aligned to the high-resolution anatomy (6 parameters). Data were 3D motion corrected (trilinear interpolation, with the first volume of the first run of each participant as reference), followed by slice time correction and high-pass filtering (cutoff frequency of 3 cycles per run). Spatial smoothing was applied with a Gaussian kernel of 8 mm FWHM for univariate analysis and 3 mm FWHM for MVPA (see also 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 (RFX) general linear model (GLM) 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). We used 10000 Monte Carlo simulations and a corrected cluster threshold of p = 0.05. 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.

Multivoxel pattern analysis (MVPA). MVPA was carried out using the CoSMoMVPA toolbox (Oosterhof et al., 2016). Design matrices contained 16 predictors reflecting the action conditions (8 actions x 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 6 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 6 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 beta maps (8 runs x 2 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, we projected the mean accuracy maps and the outlines of the corrected clusters on the same cortex-based aligned group surface. Decoding analyses were carried out using a linear discriminant analysis (LDA) 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 thus was 12.5%. For within-category MVPA, only the two actions of the same category were decoded (e.g. open vs. close for the transitive/ non-social action category). The decoding accuracy at chance was 50%. For both analyses, classification accuracies were computed using leave-one-out cross validation, i.e., 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 carried out 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, we collapsed the beta values of the two actions within each category. We used a cross decoding scheme: To decode actions along transitivity, we trained the classifier to discriminate between transitive vs. 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 vs. 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 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.

Representational similarity analysis (RSA). For RSA (Kriegeskorte et al., 2008a), we averaged the 16 beta values of each action condition for each participant and voxel. For each searchlight/ROI sphere, we extracted the mean beta values to obtain one multivoxel pattern per action. For each pattern, we normalized the beta values by subtracting the mean beta value from each individual beta value (demeaning). Next, we correlated the patterns with each other resulting in an 8 x 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 (pSTS and parahippocampal cortex, PHC) in each hemisphere based on the peak coordinates of the univariate conjunctions of sociality and transitivity (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, whose anchor points are based on univariate activation differences, 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 mid 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). 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 DLOTC and VLOTC using hierarchical cluster analysis. DLOTC and VLOTC 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 (Fig. 7) and then defined DLOTC and VLOTC ROIs dorsally and ventrally of that peak, i.e., 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 ± 1.3%, SEM).

Behavioral ratings for sociality and transitivity revealed that actions were clearly categorized into transitive vs. intransitive and social vs. 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 like 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 (Tab. 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., 2012; Tucciarelli et al., 2015; Wurm et al., 2015; Wurm and Lingnau, 2015).

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 vs. B and tested the same

385 classifier using the categories C vs. D (and vice versa). Concretely, to decode
386 sociality-specific features, we trained a classifier to decode change of possession
387 (trans/social) vs. object manipulation (trans/nonsocial) and tested the classifier using
388 communication (intrans/social) vs. body/contact actions (intrans/nonsocial). Likewise,
389 to decode transitivity-specific features, we trained a classifier to decode change of
390 possession (trans/social) vs. communication (intrans/social) and tested the classifier
391 using object manipulation (trans/nonsocial) vs. body/contact actions
392 (intrans/nonsocial).

393 Both searchlight analyses revealed strong above chance accuracies in bilateral LOTC
394 and, strikingly, far weaker effects in parietal regions (Fig. 2B). Decoding of sociality
395 and transitivity features differed mostly with respect to overall decoding strength, i.e.,
396 there were higher decoding accuracies for transitivity in comparison to sociality. This
397 difference is not surprising because transitivity distinguishes actions based on salient
398 perceptual features such as reaching and grasping of objects whereas sociality
399 distinguishes actions based on more subtle, probably less perceptual features.
400 However, there were also anatomical differences: While decoding of transitivity
401 comprised regions in dorsal and ventral LOTC as well as in VOTC, decoding of
402 sociality was mostly restricted to dorsal LOTC. Critically, in both hemispheres
403 decoding peaks of transitivity were in ventral LOTC while peaks of sociality were in
404 dorsal LOTC (Tab. 1, Fig. 5).

405 In a second step, we characterized the representational organization of brain regions
406 with respect to sociality and transitivity. To this end, we performed a searchlight-
407 based RSA using representational dissimilarity matrices (RDM) obtained from
408 behavioral ratings for sociality and transitivity (Fig. 3). The RSA for sociality
409 revealed significant effects in bilateral pMTG as well as in left postcentral gyrus. The

RSA for transitivity revealed significant effects throughout lateral and ventral OTC (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 ITG (ventral; Lingnau & Downing, 2015), the clusters found for sociality were located more dorsally than those for transitivity (Tab. 2). At a larger topographical scale, however, the dorsal-ventral gradient from transitivity to sociality was less strict as there were nearby regions fitting the transitivity model in regions other than ITG/FG/PHC (left posterior operculum/SMG and right posterior operculum/STG), 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, i.e., information that generalizes across category-specific actions. It is unclear, however, how these abstract dimensions neuroanatomically relate 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 vs. nonsocial and transitive vs. 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

435 only with respect to start and end location of the object relative to the actor (note that
436 due to the variance of actor position - left or right side of the table - and perspective, it
437 is impossible that the decoding relied on absolute object positions). Hence, in the
438 within-category MVPA, classification due to perceptual differences was minimized
439 by keeping category-specific features, such as reaching and grasping, constant. By
440 contrast, in the across-category MVPA, classification due to perceptual differences
441 was minimized by generalizing across category-specific features. In addition, for both
442 approaches, the high stimulus variance ensured that decoding relies on abstract
443 representations that generalize across features like effector, perspective, etc.

444 In a first step, we performed searchlight analyses for each category separately. For
445 each category, we obtained mean accuracy and *t*-maps to reveal regions where
446 decoding accuracy was consistently above chance (50%) across participants (Fig. 4A,
447 Tab. 3). Decoding accuracies were generally highest in occipitotemporal and parietal
448 regions; however, not all four searchlight analyses revealed statistically robust effects
449 surviving TFCE correction. Decoding *open* vs. *close* (object manipulation) revealed
450 significant clusters in left LOTC and left postcentral sulcus. These clusters overlapped
451 well with the clusters found in a previous study that decoded *open* vs *close* actions
452 using different stimuli (Wurm and Lingnau, 2015). Decoding *agree* vs. *disagree*
453 (communication) revealed similar clusters in left and right LOTC and left postcentral
454 sulcus ($p < 0.001$, uncorrected), but only the cluster in right LOTC survived TFCE
455 correction. Decoding *give* vs. *take* (change of possession) revealed a cluster in right
456 LOTC ($p < 0.005$, uncorrected). Decoding *stroke* vs. *scratch* (contact/body action)
457 revealed a cluster in right precentral gyrus/sulcus ($p < 0.0025$, uncorrected). A
458 comparison of the maps revealed no systematic segregation in LOTC along
459 transitivity and/or sociality. This is perhaps not surprising because any higher-level

460 information such as sociality and transitivity is constant between the decoded actions
461 in the within-category MVPA and was thus canceled out. However, because the
462 division into four separate searchlight analyses naturally reduced the power for each
463 analysis, we cannot rule out that also the representation of more concrete features of
464 category-specific actions reflects distinctions along transitivity and/or sociality.

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

476 Though not the focus of the current study, the results obtained in parietal regions are
477 worth mentioning. The clusters in PoCS partly overlapped with the anterior inferior
478 parietal peak of the multiclass decoding. Interestingly, anterior IPL was found only to
479 a weak extent in the across-category decoding. In line with previous findings
480 (Oosterhof et al., 2010; Oosterhof et al., 2012; Leshinskaya and Caramazza, 2015;
481 Wurm et al., 2015; Wurm and Lingnau, 2015), this pattern of results suggests that left
482 anterior IPL represents action-specific information of a high degree of generality but
483 is less likely to represent higher order dimensions like sociality and transitivity.
484 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 like sociality and transitivity. At the same time, IPS/SPL did not differentiate actions of the same category when they were perceptually very similar, e.g. have similar movement trajectories. Taken together, these findings suggest that IPS/SPL codes coarse-grained spatial action features specific for each of the categories. In line 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 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 towards the non-acting 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 persons feelings and reactions (Saxe and Kanwisher, 2003; Wurm et al., 2011). By contrast, observation of transitive actions is likely to direct attention towards 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 regions representing person and inanimate object information, respectively. To identify regions showing such activation differences we computed univariate contrast conjunctions for social vs. nonsocial (*give/take* vs. *open/close* and *agree/disagree* vs. *stroke/scratch*) and transitive vs. intransitive actions (*give/take* vs. *agree/disagree* and *open/close* vs. *stroke/scratch*), respectively (Fig. 6). The contrast conjunction social vs. nonsocial revealed bilateral posterior superior temporal sulcus (pSTS), i.e., 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 vs. 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 and SPL, i.e., 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 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 vs. nonsocial and transitive vs. 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 beta estimates and plotted these responses as a function of the position on the dorsal-ventral axis. For univariate effects, we computed beta estimates for each of the four action categories separately. For a better visualization of the relative differences between categories we normalized beta 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 Fig. 7. There are two major findings: (1) 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. (2) 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 to the ventral side of this peak. By contrast, effects of transitivity were stronger on the ventral compared to the dorsal side of this peak. These peaks were located between pSTS (dorsal end of the LOTC), the multiclass decoding peak in LOTC (mid of LOTC) and ITG (ventral end of LOTC; border to VOTC). In line 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 action categories (*give/take* and *open/close*), 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 DLOTC (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 DLOTC, social and

584 nonsocial actions formed superordinate clusters; in VLOTC, transitive and intransitive
585 actions formed superordinate clusters (Fig. 8).

586 Finally, to investigate the gradient from action-specific to more general action
587 features along the posterior-anterior axis, we conducted a second vector-of-ROI
588 analysis. For investigation of this action-specificity gradient, only the within- and
589 across-category decoding is informative. We therefore performed only these decoding
590 analyses (Fig. 9). In line with the whole-brain analysis (Fig. 5), in both hemispheres
591 the peaks of the across-category decoding were located more anteriorly relative to the
592 within-category decoding. Note however, that the within-category decoding revealed
593 only subtle variations along the posterior-anterior axis, i.e., there was no clearly
594 outstanding peak. This analysis therefore provided only moderate evidence for an
595 action-specific-to-general gradient along the posterior axis.

596

597

598 **DISCUSSION**

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

601 (1) Features associated with social vs. nonsocial and transitive vs. intransitive actions
602 could be decoded in LOTC independently of the specific action category. For
603 example, a classifier that was trained to distinguish between change of possession
604 (social/transitive) and object manipulation (non-social/transitive) actions was able to
605 distinguish between communicative (social/non-transitive) and body/contact (non-
606 social/non-transitive) actions. This finding suggests that LOTC represents features of
607 of sociality and transitivity at a level that is independent of specific action subtypes.

(2) Dorsal and ventral subregions of LOTC were preferentially organized 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.

(3) 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 vs. nonsocial and transitive vs. intransitive action features, respectively. Using RSA, we demonstrated that DLOTC preferentially represents actions as predicted by the sociality model whereas VLOTC preferentially represents actions as predicted by the transitivity model. In addition, in both hemispheres the peak location of the social vs. nonsocial action decoding was dorsal to the peak location of the transitive vs. 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 (Fig. 2, Fig. 7). By 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 like those in our study – have structural

633 similarities (typically involve the dynamic processing of motion and change, are
634 typically intentional, etc.) and are therefore likely to be represented by neural
635 substrates with similar computational properties (Kaas and Catania, 2002; Rosa and
636 Tweedale, 2005). In other words, actions like *open*, *give*, *agree*, and *scratch* are more
637 similar to each other than to other, structurally different kinds of information like
638 persons and inanimate objects, even if these kinds of information are important (albeit
639 not constitutive) for action recognition. On this reasoning, our finding that action
640 information was encoded in proximity but non-overlapping with person-related and
641 inanimate object knowledge, is plausible. The subdivision within action-processing
642 neural substrates along the dorsal-ventral axis, i.e., DLOTC is more sensitive to
643 sociality features whereas VLOTC is more sensitive to transitivity features, can be
644 explained under the assumption that the neuroanatomical organization of action
645 knowledge is shaped by systematic connections between object and action
646 representations: Socially relevant person information in dorsal areas such as the STS
647 should be more strongly connected to social action representations in LOTC. By
648 contrast, inanimate object information in ventral areas such as the ITG and FG should
649 be more strongly connected to object-directed action representations in LOTC. The
650 connections to person-related and inanimate object information thus might exert
651 opposing constraints on the representations of social and transitive actions, which
652 could drive the anatomical segregation in the observed way. This interpretation is
653 supported by recent studies that demonstrated enhanced functional connectivity
654 specific for inanimate objects (artifacts and tools) between FG and a region in LOTC
655 overlapping with the region we found to be sensitive for transitive vs. intransitive
656 action discrimination (Hutchison et al., 2014; Stevens et al., 2015). Likewise,

657 effective connectivity between pSTS and LOTC is modulated by socially relevant
658 cues like facial expressions (Furl et al., 2015).

659 What remains unspecified is the kind of information that drives the observed
660 distinctions in DLOTC and VLOTC as revealed by the RSA and the across-category
661 decoding. Do the distinctions reflect semantic categorizations or are they driven by
662 structural properties of the observed actions? Transitive actions can indeed be
663 differentiated from intransitive actions based on intrinsic structural properties such as
664 the reaching and grasping of objects. It is reasonable to assume that neural systems
665 important for the recognition of reaching and grasping as well as hand-object
666 interaction would be located in proximity to regions coding tools and other graspable
667 objects (Bracci et al., 2012). The high structural similarities of actions within the
668 transitive and intransitive categories are also reflected in the overall higher accuracies
669 of the across-category decoding for transitivity. For social actions, perceptual
670 commonalities are less evident. *Give* and *take* are perceptually different from *agree*
671 and *disagree* gestures, and likewise *open* and *close* are perceptually different from
672 *stroke* and *scratch* actions. In line with this view, multivariate effects of sociality were
673 generally subtler than effects of transitivity. Furthermore, in both social and nonsocial
674 actions an attentive passive person was present ruling out that social actions could be
675 distinguished from nonsocial actions based on perceptual cues of the passive person.
676 Increased processing of the passive person for social actions is unlikely to drive the
677 distinction because in that case one should also have observed univariate activation
678 differences between social and nonsocial actions in LOTC, which was not the case.
679 However, the social actions were directed to another person and can thus be
680 interpreted as interpersonal actions, even if there was no observable reaction of the
681 passive person. For the social actions, the acting and the passive person therefore

682 defined a common social space, which was less the case for the nonsocial actions. The
683 distinction between social and nonsocial could therefore be explained by more general
684 underlying dimensions such as social space or the direction of an action toward
685 another person or not. Another possibility is that general differences in the complexity
686 of fine hand/finger movements, independent of the concrete movements themselves,
687 drove the distinction between social vs. nonsocial actions. Indeed, we found stronger
688 univariate responses for the nonsocial vs. social actions at the level of the pMTG (Fig.
689 7). Note however that the univariate response profile of the nonsocial actions differed
690 from the profile of the sociality RSA, which suggests that the two analyses picked up
691 different kinds of information. Finally, it is possible that the across-category decoding
692 relied on semantic representations of action primitives (Schank, 1973; Schank and
693 Abelson, 1977), that were similar for the social actions and the transitive actions,
694 respectively. In fact, the social actions used in our study (give, take, agree, disagree)
695 involved a *transfer* of (physical or mental) objects. At the same time, the transitive
696 actions (give, take, open, close) involved a *change* (of location or configuration) of
697 objects. Action concepts that are composed of similar action primitives would
698 therefore be close to each other in representational space, in line with our findings.
699 Future studies should investigate the extent to which such decompositional models
700 (Jackendoff, 1972; Gruber, 1976; Pinker, 1989) can explain the neural organization of
701 action knowledge.

702

703 **Posterior to anterior LOTC is organized along a gradient from action-specific to**
704 **general action information.** A secondary finding of our study is that cluster peaks of
705 the within-category decoding were located more posteriorly in LOTC than the peaks
706 of the across-category decoding. Note however, that the range of accuracies of the

707 within-category decoding was relatively shallow and the clusters of both decoding
708 analyses showed substantial overlaps. The analysis therefore suggests only subtle,
709 preferential differences of representational content along the posterior-anterior axis.

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

723 By contrast, the across-category decoding was not suited to detect information
724 specific for action subtypes because the classifier was trained and tested on actions of
725 different categories (e.g., trained on change of possession vs. object manipulation and
726 tested on communication vs. body/contact actions). As elaborated above, the across-
727 category decoding was most sensitive to action features that generalize across
728 categories along the dimensions sociality and transitivity. Taken together, the
729 different peak locations of within- and across-category decoding suggest that abstract
730 action-general features are represented more anteriorly than concrete action-specific
731 features, which is in line with recent proposals on the functional organization of

LOTC from concrete to abstract and from visual to amodal action representations (Watson and Chatterjee, 2011; Lingnau and Downing, 2015; see also Thompson-Schill, 2003; Martin, 2007).

Conclusions

Our results suggest a topographic organization of LOTC along two major axes: a dorsal vs. ventral distinction that segregates social vs. 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 vs. artifacts, and their connectivity. Together, our results help establishing 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.

Allison T, Puce A, McCarthy G (2000) Social perception from visual cues: role of the STS region. Trends Cogn Sci 4:267-278.

Beauchamp MS, Lee KE, Haxby JV, Martin A (2002) Parallel visual motion processing streams for manipulable objects and human movements. Neuron 34:149-159.

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.

Binkofski F, Buxbaum LJ (2013) Two action systems in the human brain. Brain Lang 127:222-229.

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.

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.

Brainard DH (1997) The Psychophysics Toolbox. Spatial vision 10:433-436.

Caramazza A, Shelton JR (1998) Domain-specific knowledge systems in the brain the animate-inanimate distinction. J Cogn Neurosci 10:1-34.

Caramazza A, Mahon BZ (2003) The organization of conceptual knowledge: the evidence from category-specific semantic deficits. Trends Cogn Sci 7:354-361.

Carter RM, Huettel SA (2013) A nexus model of the temporal-parietal junction. Trends Cogn Sci 17:328-336.

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.

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.

Downing PE, Chan AW, Peelen MV, Dodds CM, Kanwisher N (2006) Domain specificity in visual cortex. Cereb Cortex 16:1453-1461.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. Neuroimage 7:30-40.

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.

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.

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.

Jackendoff RS (1972) Semantic Interpretation in Generative Grammar. Cambridge, Mass: MIT Press.

797 Kaas JH, Catania KC (2002) How do features of sensory representations develop?
798 BioEssays : news and reviews in molecular, cellular and developmental
799 biology 24:334-343.
800 Konkle T, Caramazza A (2013) Tripartite organization of the ventral stream by
801 animacy and object size. J Neurosci 33:10235-10242.
802 Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain
803 mapping. Proc Natl Acad Sci U S A 103:3863-3868.
804 Kriegeskorte N, Mur M, Bandettini P (2008a) Representational similarity analysis -
805 connecting the branches of systems neuroscience. Frontiers in systems
806 neuroscience 2:4.
807 Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K,
808 Bandettini PA (2008b) Matching categorical object representations in inferior
809 temporal cortex of man and monkey. Neuron 60:1126-1141.
810 Leshinskaya A, Caramazza A (2015) Abstract categories of functions in anterior
811 parietal lobe. Neuropsychologia.
812 Lingnau A, Downing PE (2015) The lateral occipitotemporal cortex in action. Trends
813 Cogn Sci 19:268-277.
814 Mahon BZ, Milleville SC, Negri GA, Rumiati RI, Caramazza A, Martin A (2007)
815 Action-related properties shape object representations in the ventral stream.
816 Neuron 55:507-520.
817 Martin A (2007) The representation of object concepts in the brain. Annu Rev
818 Psychol 58:25-45.
819 Nichols T, Brett M, Andersson J, Wager T, Poline JB (2005) Valid conjunction
820 inference with the minimum statistic. Neuroimage 25:653-660.
821 Oosterhof NN, Tipper SP, Downing PE (2012) Viewpoint (in)dependence of action
822 representations: an MVPA study. J Cogn Neurosci 24:975-989.
823 Oosterhof NN, Connolly AC, Haxby JV (2016) CoSMoMVPA: Multi-Modal
824 Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave.
825 Front Neuroinform 10:27.
826 Oosterhof NN, Wiggett AJ, Diedrichsen J, Tipper SP, Downing PE (2010) Surface-
827 based information mapping reveals crossmodal vision-action representations
828 in human parietal and occipitotemporal cortex. J Neurophysiol 104:1077-
829 1089.
830 Orlov T, Porat Y, Makin TR, Zohary E (2014) Hands in motion: an upper-limb-
831 selective area in the occipitotemporal cortex shows sensitivity to viewed hand
832 kinematics. J Neurosci 34:4882-4895.
833 Pillon A, d'Honincthun P (2011) The organization of the conceptual system: the case
834 of the "object versus action" dimension. Cognitive neuropsychology 27:587-
835 613.
836 Pinker S (1989) Learnability and Cognition: The Acquisition of Argument Structure.
837 Cambridge: The MIT Press.
838 Rosa MG, Tweedale R (2005) Brain maps, great and small: lessons from comparative
839 studies of primate visual cortical organization. Philos Trans R Soc Lond B
840 Biol Sci 360:665-691.
841 Saxe R, Kanwisher N (2003) People thinking about thinking people. The role of the
842 temporo-parietal junction in "theory of mind". Neuroimage 19:1835-1842.
843 Schank RC (1973) The fourteen primitive actions and their inferences. Memo AIM-
844 183, Stanford Artificial Intelligence Laboratory.
845 Schank RC, Abelson RP (1977) Scripts, plans, goals, and understanding: An inquiry
846 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. Behavior research methods 43:1194-1201.
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.
Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44:83-98.
Stevens WD, Tessler MH, Peng CS, Martin A (2015) Functional connectivity constrains the category-related organization of human ventral occipitotemporal cortex. Hum Brain Mapp 36:2187-2206.
Thompson-Schill SL (2003) Neuroimaging studies of semantic memory: inferring "how" from "where". Neuropsychologia 41:280-292.
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.
Turella L, Lingnau A (2014) Neural correlates of grasping. Front Hum Neurosci 8:686.
Van Overwalle F, Baetens K (2009) Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. Neuroimage 48:564-584.
Watson CE, Chatterjee A (2011) The functional neuroanatomy of actions. Neurology 76:1428-1434.
Wurm MF, Lingnau A (2015) Decoding actions at different levels of abstraction. J Neurosci 35:7727-7735.
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.
Wurm MF, Ariani G, Greenlee MW, Lingnau A (2015) Decoding Concrete and Abstract Action Representations During Explicit and Implicit Conceptual Processing. Cereb Cortex.
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. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 52:1156-1166.
Zeng H, Constable RT (2002) Image distortion correction in EPI: comparison of field mapping with point spread function mapping. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 48:137-146.

TABLES

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

Region	x	y	z	*t*	*p*	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

Peak coordinates of corrected clusters in Talairach coordinates (x,y,z). Abbreviations: aIPS, anterior intraparietal sulcus; FG, fusiform gyrus; IFG, inferior frontal gyrus; IPS, intraparietal sulcus; LOTC, lateral occipitotemporal cortex; PHC, parahippocampal cortex; PMd, dorsal premotor cortex; PMv, ventral premotor cortex; PoCS, postcentral sulcus; SMG, supramarginal gyrus; SPL, superior parietal lobe, TOS, transverse occipital sulcus.

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

Region	x	y	z	*t*	*p*
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

Peak coordinates of corrected clusters in Talairach coordinates (x,y,z). Abbreviations: FG, fusiform gyrus; intraparietal sulcus; LOTC, lateral occipitotemporal cortex; PHC, parahippocampal cortex; PoCG, postcentral gyrus; SMG, supramarginal gyrus; SPL, superior parietal lobe, VOTV, ventral occipitotemporal cortex.

Table 3. Clusters identified in within-category decoding

Region	x	y	z	*t*	*p*	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 vs. 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 vs. 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 vs. 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 vs. scratch						
right PMd	45	-4	43	3.79	7.37E-04*	58.1

Peak coordinates of uncorrected (indicated by Asterisks) and corrected clusters in Talairach coordinates (x,y,z). Abbreviations: aIPS, anterior intraparietal sulcus; IPS, intraparietal sulcus; LOTC, lateral occipitotemporal cortex; MTG, middle temporal gyrus; PMd, dorsal premotor cortex; PoCS, postcentral sulcus; SMG, supramarginal gyrus; SPL, superior parietal lobe.

914 **Table 4.** Clusters identified in univariate contrast conjunctions

Region	x	y	z	*t*	*p*
social vs. nonsocial					
left pSTS	-46	-52	20	4.13	2.93E-04
right pSTS	47	-63	23	3.91	5.38E-04
transitive vs. 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

915

916 Peak coordinates of corrected clusters in Talairach coordinates (x,y,z). Abbreviations: FG, fusiform gyrus; PHC,
917 parahippocampal cortex; PMd, dorsal premotor cortex; pSTS, posterior superior temporal sulcus; SPL, superior
918 parietal lobe.

919

FIGURE CAPTIONS

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) Representational dissimilarity matrices (RDMs) for sociality and transitivity obtained from behavioral ratings (averaged across participants).

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.

Figure 3. Representational similarity analysis. 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.

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.

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 9 adjacent vertices with the next highest t values in an iterative manner, i.e., 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; ITG, inferior temporal gyrus; ITS, inferior temporal sulcus; LnS, lunate sulcus; LOG, lateral occipital gyrus; LOS, lateral occipital sulcus; MTG, middle temporal gyrus; STS, superior temporal 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.

Figure 7. Vector-of-ROI analysis along the dorsal-ventral axis. To investigate the sociality-transitivity gradient, decoding accuracies (multiclass, across- and within-category decoding of transitivity and sociality), RSA correlations of neural RDMs with the transitivity and sociality models, and univariate beta estimates for each action category are plotted as a function of position on the dorsal-ventral axis from pSTS to PHC (see 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

969 sociality and transitivity, respectively, as revealed by the across-category decoding
970 and the RSA.

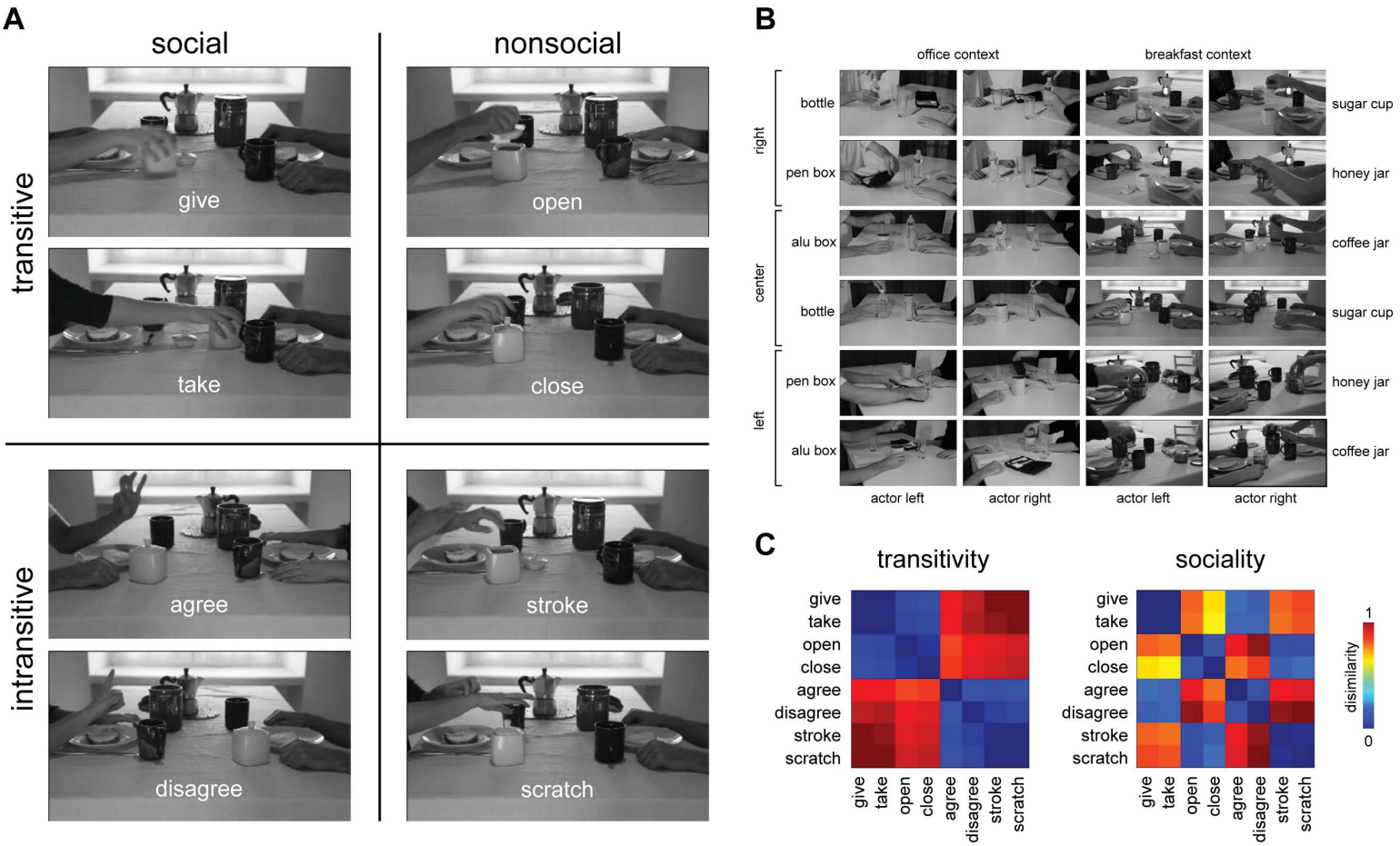
971

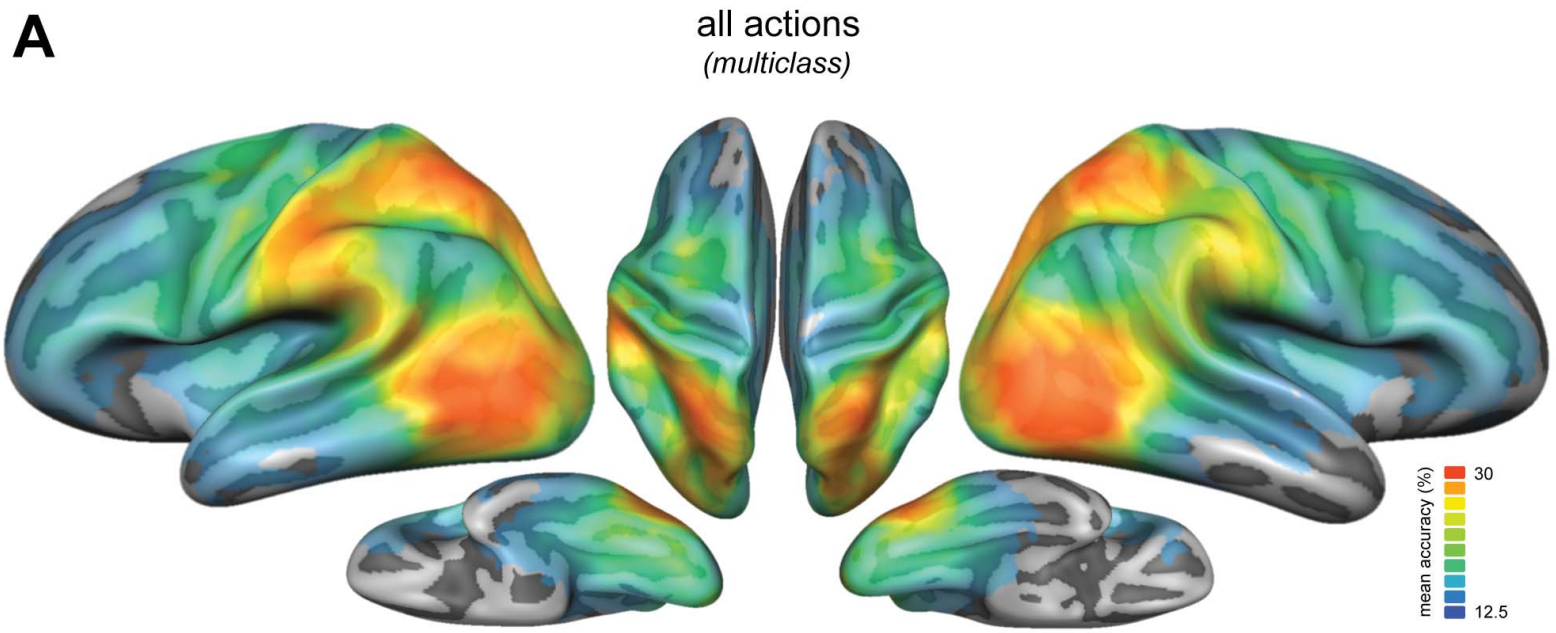
972 **Figure 8.** Dendrogram plots of the hierarchical cluster analysis (average distance) for
973 dorsal and ventral LOTC. ROIs were defined by selecting eight adjacent ROIs of the
974 vector-of-ROI analysis located dorsally (DLOTC) and ventrally (VLOTC) of the peak
975 ROI of the multiclass decoding (see Fig. 7).

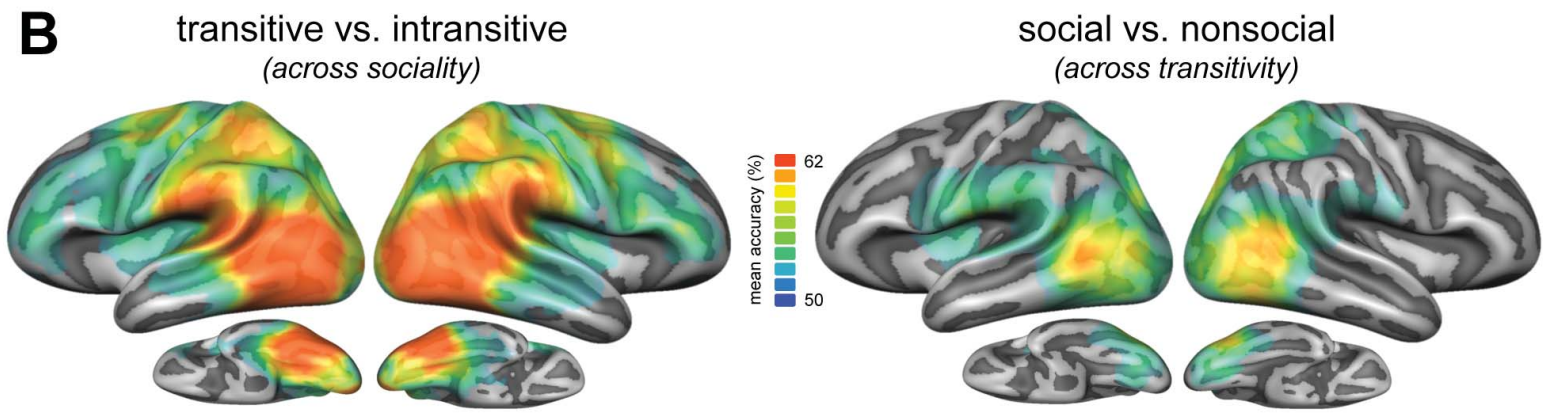
976

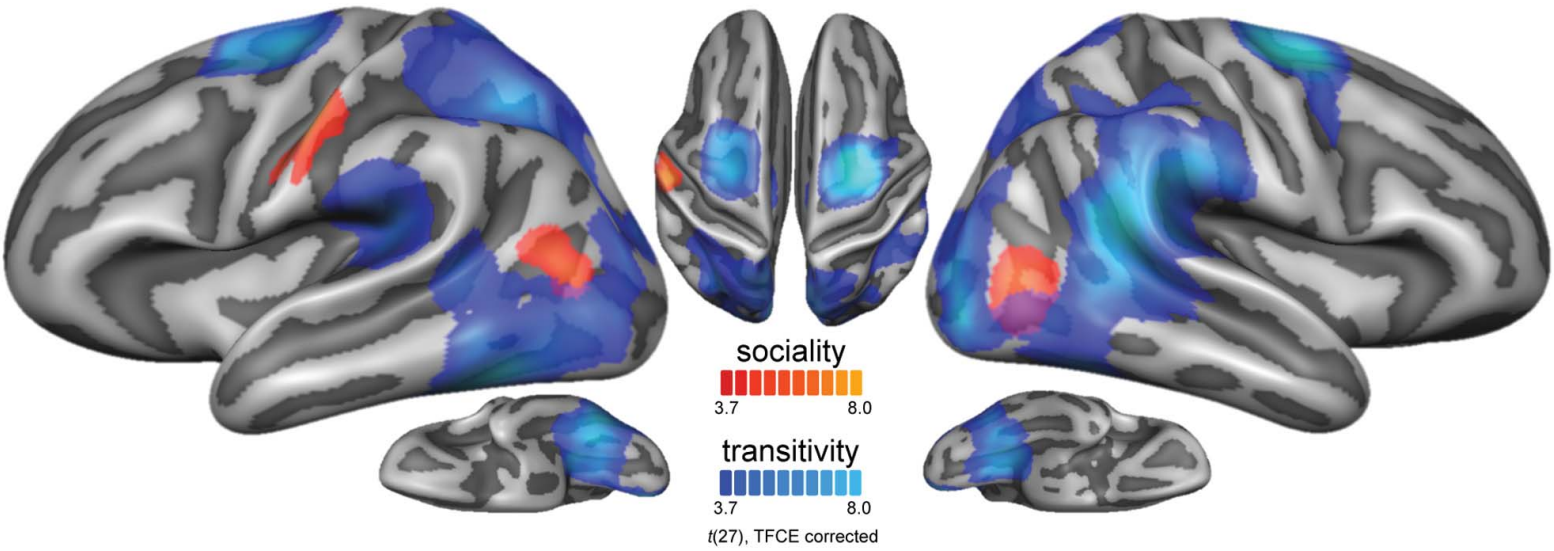
977 **Figure 9.** Vector-of-ROI analysis along the posterior-anterior axis. To investigate the
978 gradient from action-specific to general action information, across- and within-
979 category decoding accuracies are plotted as a function of position on the posterior-
980 anterior axis from early visual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

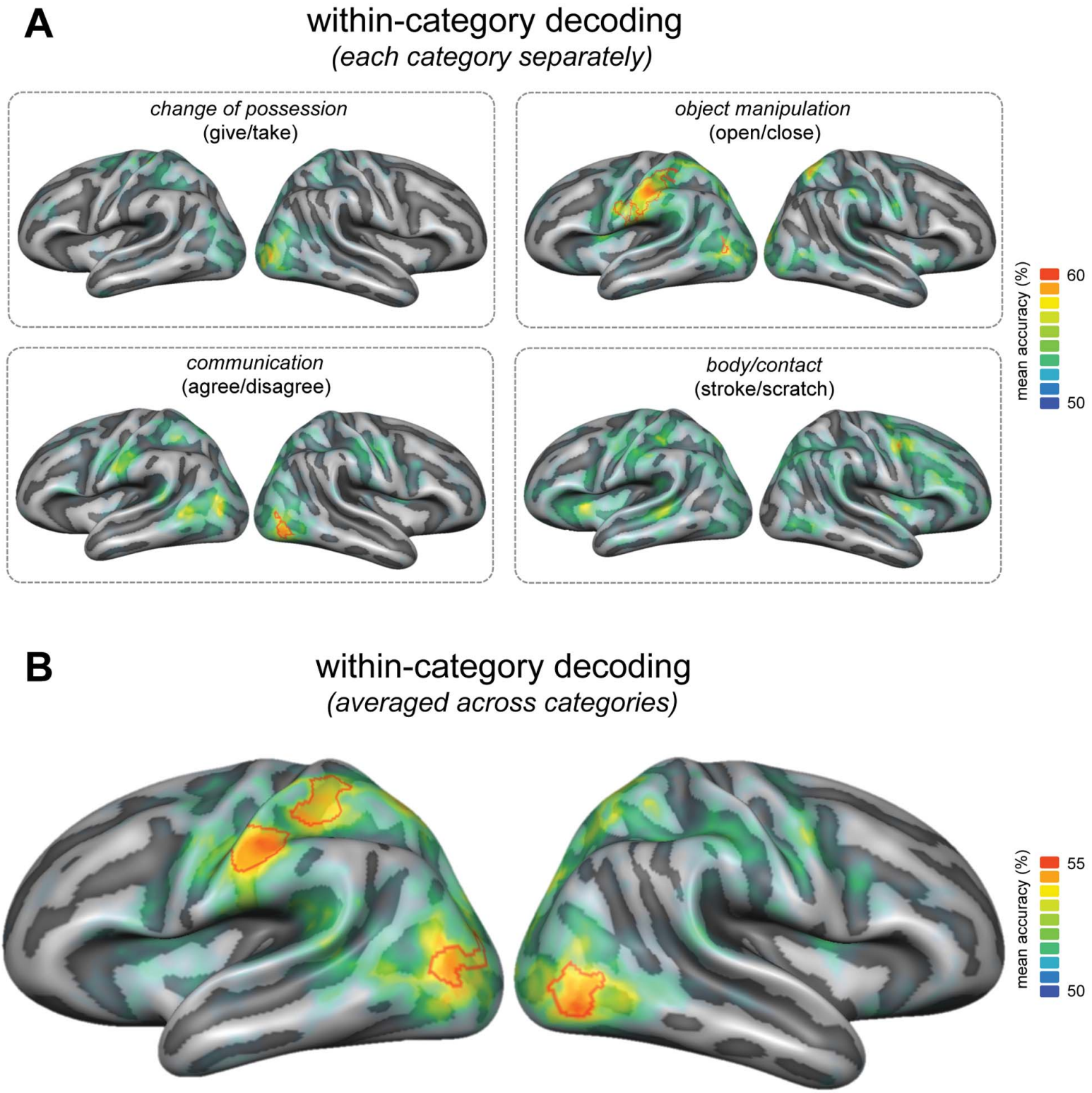
984

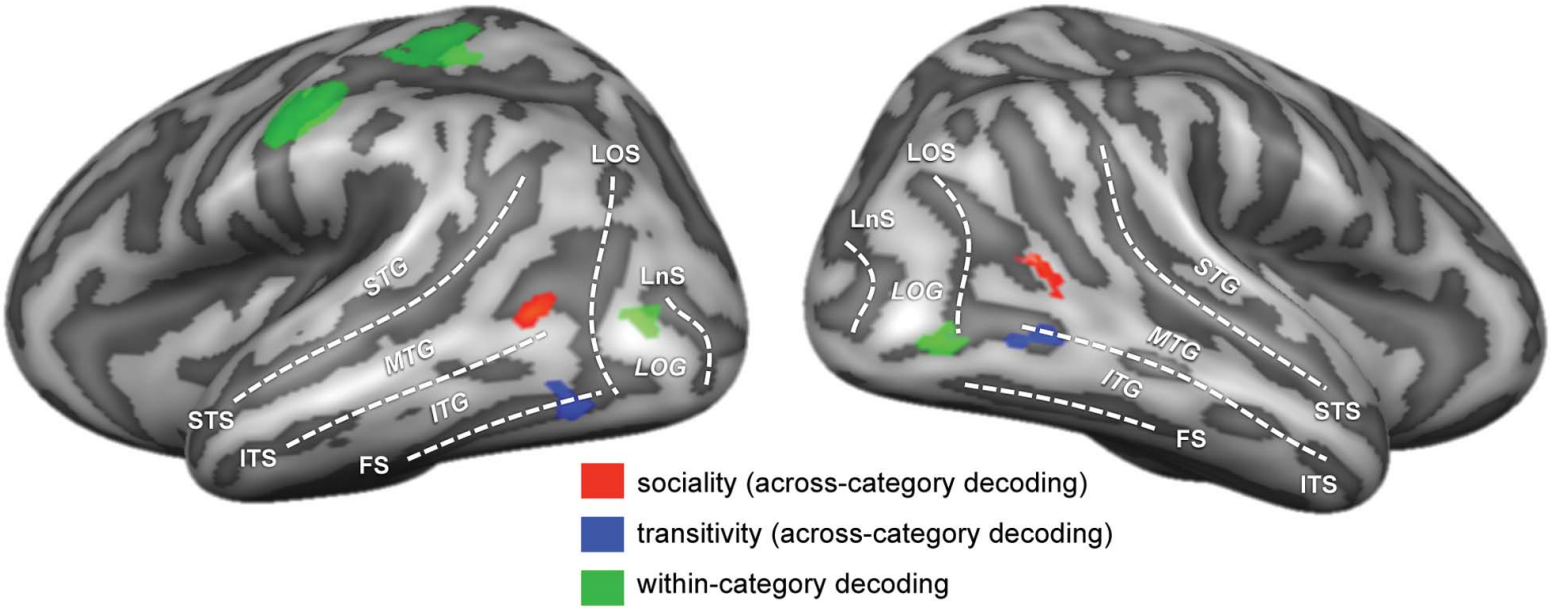

A
social
nonsocial
transitive
intransitive
give
open
take
close
agree
stroke
disagree
scratch
B
office context
breakfast context
right
center
left
bottle
pen box
alu box
bottle
pen box
alu box
sugar cup
honey jar
coffee jar
sugar cup
honey jar
coffee jar
actor left
actor right
actor left
actor right
C
transitivity
sociality
give
take
open
close
agree
disagree
stroke
scratch
dissimilarity
1
0

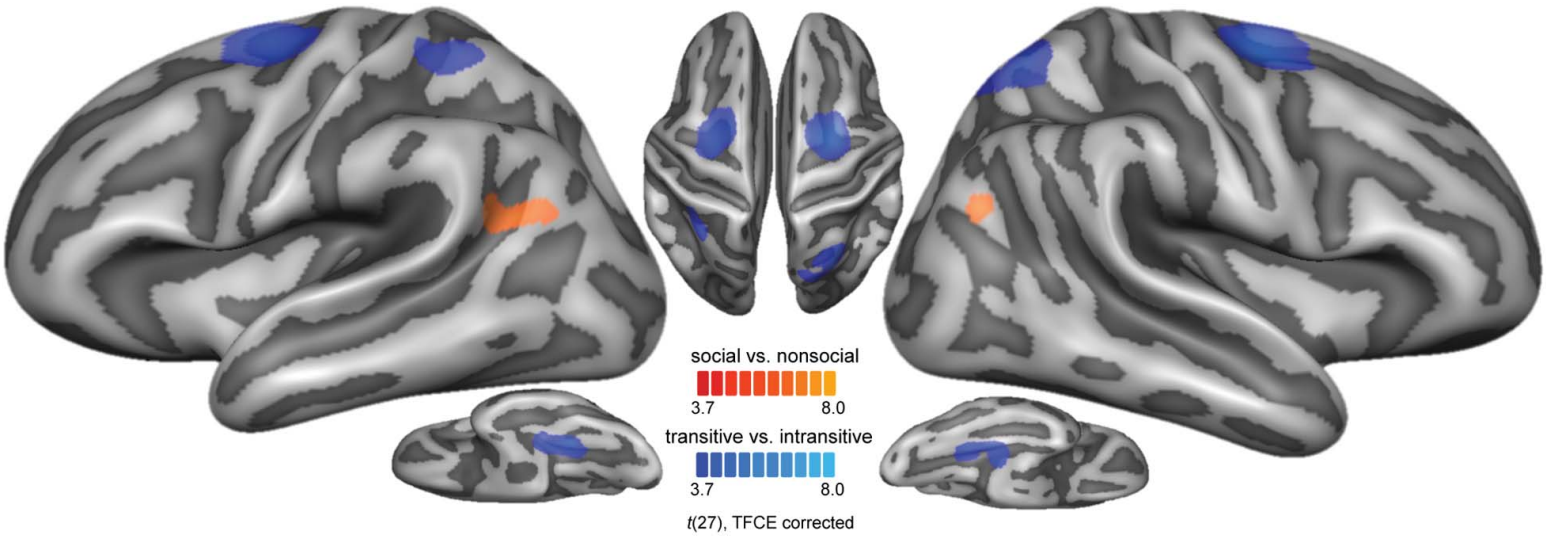

A
all actions
(multiclass)
mean accuracy (%)
30
12.5

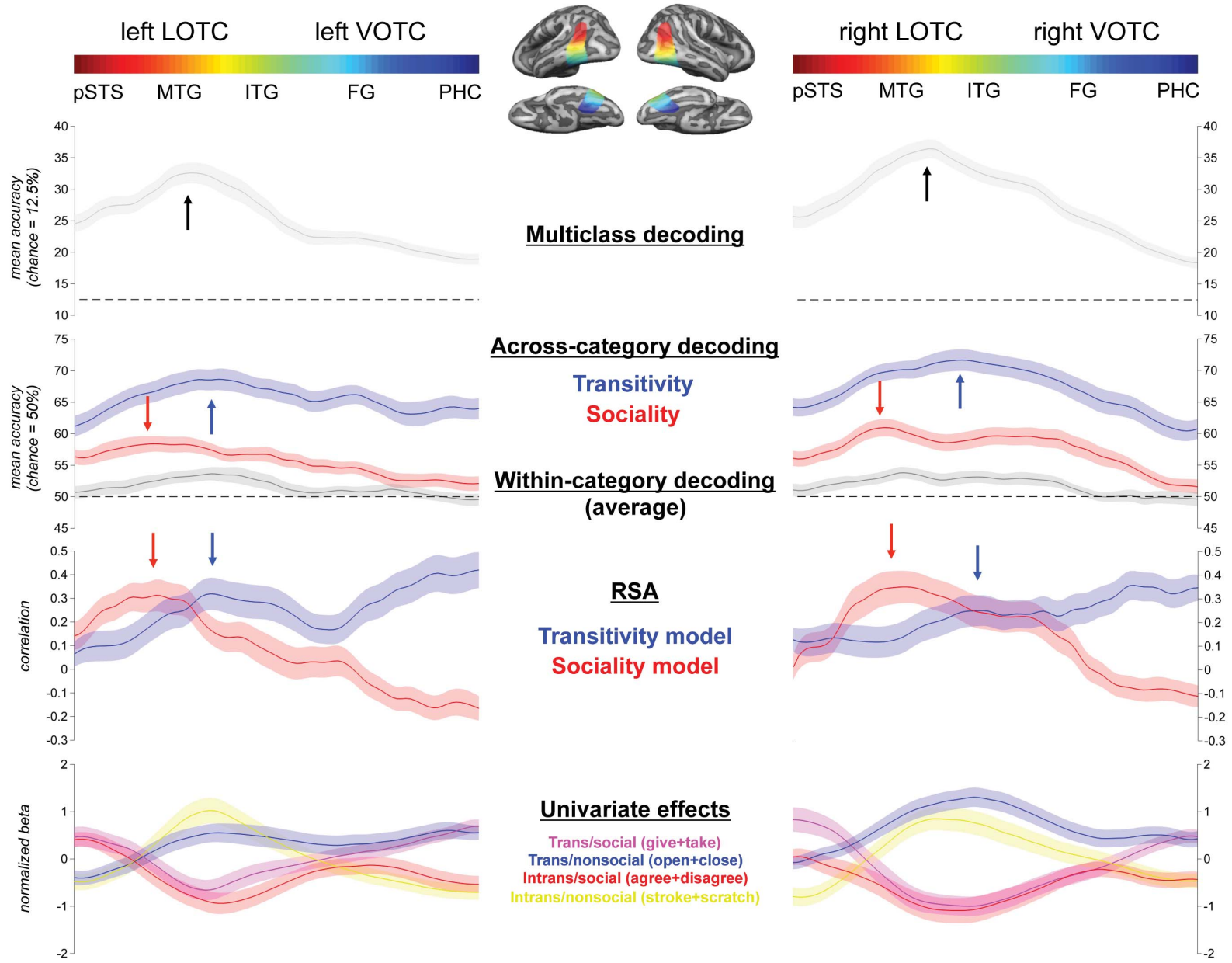

B
transitive vs. intransitive
(across sociality)
social vs. nonsocial
(across transitivity)
mean accuracy (%)
62
50

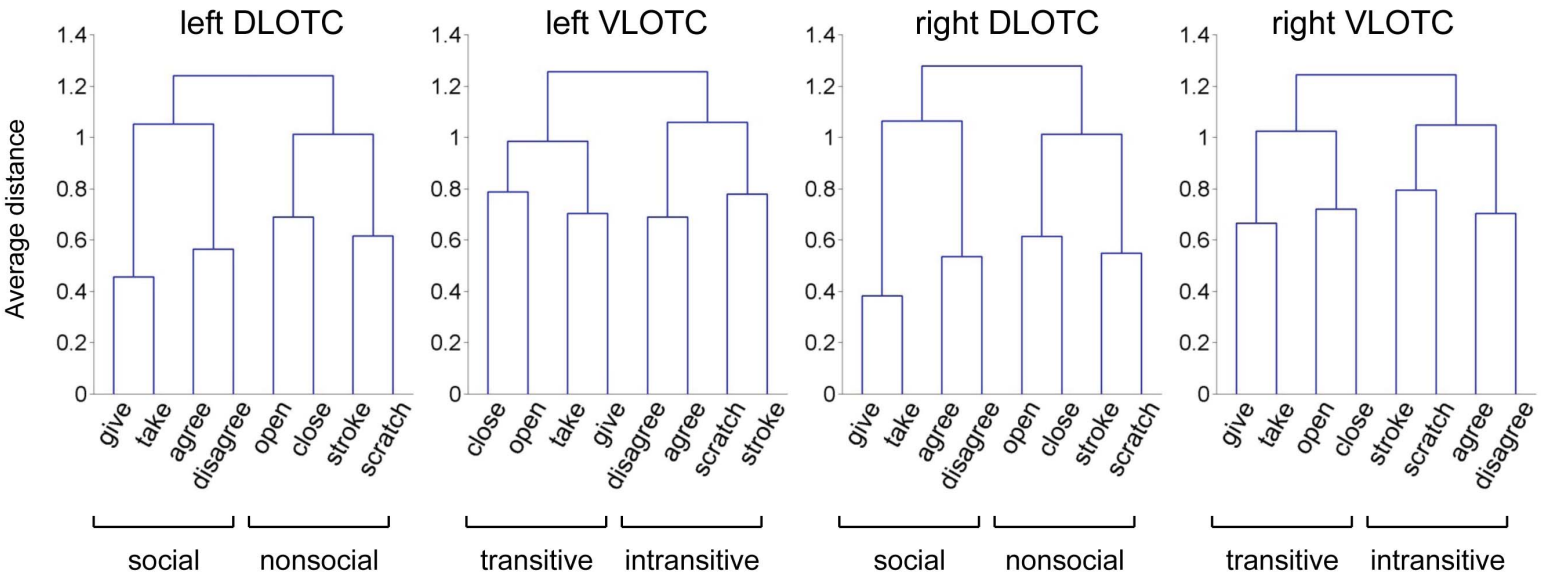

sociality
3.7
8.0
transitivity
3.7
8.0
t(27), TFCE corrected

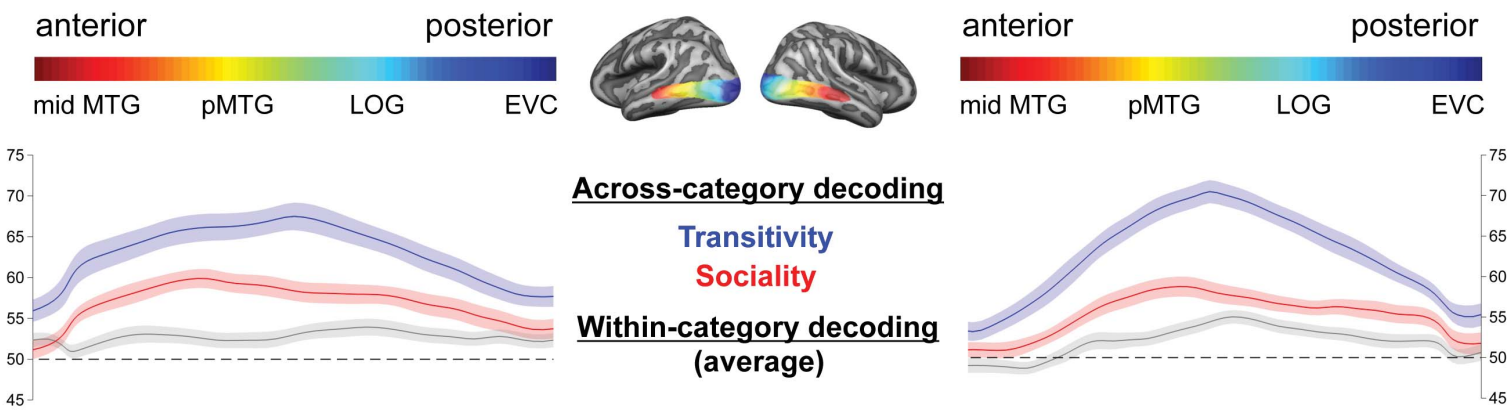

A
within-category decoding
(each category separately)
change of possession
(give/take)
object manipulation
(open/close)
communication
(agree/disagree)
body/contact
(stroke/scratch)
mean accuracy (%)
60
50
B
within-category decoding
(averaged across categories)
mean accuracy (%)
55
50


LOS
LnS
STG
MTG
ITG
LOG
STS
ITS
FS
sociality (across-category decoding)
transitivity (across-category decoding)
within-category decoding


social vs. nonsocial
3.7
8.0
transitive vs. intransitive
3.7
8.0
t(27), TFCE corrected


left LOTC
left VOTC
pSTS
MTG
ITG
FG
PHC
right LOTC
right VOTC
pSTS
MTG
ITG
FG
PHC
mean accuracy (chance = 12.5%)
40
35
30
25
20
15
10
Multiclass decoding
Across-category decoding
Transitivity
Sociality
mean accuracy (chance = 50%)
75
70
65
60
55
50
45
Within-category decoding (average)
RSA
Transitivity model
Sociality model
correlation
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
Univariate effects
Trans/social (give+take)
Trans/nonsocial (open+close)
Intrans/social (agree+disagree)
Intrans/nonsocial (stroke+scratch)
normalized beta
2
1
0
-1
-2


left DLOTC
left VLOTC
right DLOTC
right VLOTC
Average distance
1.4
1.2
1
0.8
0.6
0.4
0.2
0
give
take
agree
disagree
open
close
stroke
scratch
social
nonsocial
transitive
intransitive


anterior
posterior
mid MTG
pMTG
LOG
EVC
anterior
posterior
mid MTG
pMTG
LOG
EVC
75
70
65
60
55
50
45
Across-category decoding
Transitivity
Sociality
Within-category decoding
(average)
75
70
65
60
55
50
45