

The organization of action representations in lateral occipitotemporal cortex

Franziska Pfannerstill



Thesis submitted for the degree of
Doctor of Philosophy
Royal Holloway, University of London
Department of Psychology
September 2019

Declaration of Authorship

I, Franziska Pfannerstill, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:

Date:

Abstract

Neural representations for various types of information related to action are located in the left lateral occipitotemporal cortex (LOTc). Existing literature on basic and biological motion, bodies, body parts, tools, action perception and performance, but also action verbs demonstrated recruitment of the LOTc. This thesis investigates the architecture of this cortical area in more detail. To this aim, we carried out a direct comparison of different tasks within the same participants to examine how the different domains are represented in the LOTc in terms of their spatial and representational organization, and to relate this to the connectivity-based organization of the LOTc.

We collected functional magnetic resonance imaging (fMRI) data from $N = 21$ participants for eight different localizer experiments (motion, biological motion, bodies & tools, body parts, action observation, action performance, verbs and high level retinotopic mapping) as well as resting-state functional magnetic resonance imaging data.

First, using univariate analyses, we found that responses to the different domains were indeed located in the LOTc. Responses were arranged topographically, and subregions of LOTc were sparsely selective for only one domain. Subsequent multivariate analyses indicated a two-part organization of the LOTc where posterior-central regions show sensitivity to most contrasts and highly similar representations for body-related domains as well as spatial location information. By contrast, anterior regions were sensitive to verbs, action observation and performance, but showed no shared representations or spatial location information. Confirming the posterior-central and anterior subregion division, resting state functional connectivity analysis also indicated similar connectivity for posterior-central regions and different connections for anterior regions. The overlapping and shared representations in the posterior-central LOTc

might indicate a hybrid map of superimposed feature representations enabling action understanding.

Acknowledgements

Above all, I would like to thank my supervisor, Angelika Lingnau for the continuous support and guidance through every step of this degree. I have learned so much from you over the last three years. I am also grateful for the help, suggestions and valuable advice from Paul Downing, Jens Schwarzbach, Petra Vetter and Szonya Durant.

I am also deeply grateful for my fellow PhD colleagues and the friends I made along this journey. You all made this a lot more inspiring, bearable and enjoyable! Thank you Hanna, for being a fantastic and inspirational scientist and friend, David for always being there to joke and rant on our way home, Rachel for your kindness – go PS2010 dream-team, Adam, Clare and Jen for completing the best office corner and to Louisa for being a great bouldering companion. Special thanks also go to my housemates, Lena, Jasmine and Amanda, for always being there to laugh, worry or drink with me and making our house feel like a home.

For the immense emotional and personal support, I also thank Arno, you always encourage me, believe in me and make me laugh, thank you for being amazing.

Finally, I could not have done this without my family. Thank you for your unconditional support and love over the years, for enabling and encouraging me to pursue my goals. I am eternally grateful to you.

Content

Declaration of Authorship.....	2
Abstract.....	3
Acknowledgements.....	5
Content	6
List of Figures	8
List of Tables	12
1 Introduction	14
1.1 A hub for action perception in the brain	14
1.2 Definition of the LOTC.....	15
1.3 Activation of LOTC - Overview of diverse tasks activating LOTC	17
1.3.1 Basic motion perception	18
1.3.2 Biological motion perception.....	20
1.3.3 Body perception.....	21
1.3.4 Body parts	22
1.3.5 Tools.....	24
1.3.6 Action observation.....	25
1.3.7 Planning & Executing actions.....	27
1.3.8 Verbs	28
1.4 The hub in lateral occipitotemporal cortex	30
1.5 Organizational principles	31
1.5.1 Modules	32
1.5.2 Hierarchy.....	33
1.5.3 Topography	34
1.5.4 Distributed Processing	37
1.5.5 Networks.....	39
1.6 Architecture of the LOTC	43
2 General methods	46
2.1 Participants	46
2.2 Experimental tasks.....	46
2.2.1 Basic motion localizer	49
2.2.2 Biological motion localizer	49
2.2.3 Action execution localizer	50

2.2.4	Bodies & Tools localizer	51
2.2.5	Body parts localizer	52
2.2.6	Action observation localizer.....	53
2.2.7	Verb localizer.....	54
2.2.8	Spatial information localizer (High Level Retinotopy)	55
2.3	Procedure.....	56
2.4	Data acquisition	57
2.5	Preprocessing for functional data.....	57
2.6	ROI definition	58
3	Spatial organization in LOTC	60
3.1	Introduction	60
3.2	Analyses	60
3.2.1	Neural responses for different domains in the brain	60
3.2.2	Topography of peak activation	62
3.2.3	Selectivity for domains in LOTC subparcels	63
3.3	Results.....	63
3.3.1	Neural responses for different domains in the brain	63
3.3.2	Topography of peak activation	71
3.3.3	Selectivity of subareas	73
3.4	Discussion.....	78
4	Representational organization in LOTC	85
4.1	Introduction	85
4.2	Analyses	86
4.2.1	Sensitivity, discriminatory property and similarity for domains in LOTC subparcels 86	
4.2.2	Similarity of representations across LOTC subparcels.....	90
4.2.3	HLR in LOTC-Subregions with representational similarity analysis.....	90
4.3	Results.....	93
4.3.1	Sensitivity, discriminatory property and similarity for contrasts in LOTC subparcels.....	93
4.3.2	Similarity of representations across LOTC subparcels	104
4.3.3	HLR in LOTC-Subregions with representational similarity analysis.....	107
4.3.4	HLR across regions	107
4.3.5	Representational content evaluated by models.....	110

4.4	Discussion.....	112
5	Connectivity based organization in LOTC	116
5.1	Introduction	116
5.2	Resting state data analysis.....	117
5.2.1	Preprocessing.....	117
5.2.2	Resting state connectivity analysis	118
5.3	Results.....	119
5.4	Discussion.....	124
6	General Discussion.....	127
6.1	Chapter summary.....	127
6.2	Future directions.....	135
6.3	Conclusion.....	137
7	References	137
8	Appendices.....	151

List of Figures

Figure 1.	Approximate structural outline of the LOTC;.....	16
Figure 2.	z-map for neurosynth meta-analysis of the term ‘motion’	19
Figure 3.	z-map for neurosynth meta-analysis of the term ‘body’	22
Figure 4.	z-map for a combination of neurosynth meta-analyses of the terms ‘face’, ‘hand’, ‘foot’ and each term separately.	23
Figure 5.	z-map for neurosynth meta-analysis of the term ‘tools’	25
Figure 6.	z-map for neurosynth meta-analysis of the term ‘action observation’	26
Figure 7.	z-map for neurosynth meta-analysis of the term ‘movement’	28
Figure 8.	z-map for neurosynth meta-analysis of the term ‘verbs’	29
Figure 9.	Combination of binarized z-maps from neurosynth meta-analysis of the terms ‘motion’, ‘body’, ‘face’, ‘hand’, ‘foot’, ‘tool’, ‘action observation’, ‘movement’, ‘verb’ showing overlap of the maps.	31

Figure 10. Location of main regions involved in networks related to domains activating the LOTC.
IFG=Inferior frontal gyrus; vPMC = ventral premotor cortex; dlPMC= dorsolateral premotor cortex; SPL= superior parietal lobe; IPL= inferior parietal lobe; STC= superior temporal cortex; LOTC = lateral occipitotemporal cortex; ITC= inferior temporal cortex; FC = fusiform cortex..... 40

Figure 11. Example trials for basic motion localizer as shown inside the scanner right before task 49

Figure 12. Example trials for biological motion localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed to detect and indicate with a button press. 50

Figure 13. Example trials for movement execution localizer as shown inside the scanner right before the task..... 51

Figure 14. Example trials for Body & Tool localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed to detect and indicate with a button press while crossed out arrows show false repetitions to non-identical stimuli..... 52

Figure 15. Example trials for body part localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed to detect and indicate with a button press while crossed out arrows show false repetitions to non-identical stimuli..... 53

Figure 16. Example trials for action observation localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed 54

Figure 17. Example trials for Noun Verb localizer as shown inside the scanner right before the task depicting trials and assignment of responses to buttons/fingers..... 55

Figure 18. Example trials for High Level Retinotopy as shown inside the scanner right before task 56

Figure 20. Left lateral view of results of the univariate group level GLMX-analysis (with $p < 0.05$, Threshold free cluster enhancement (TFCE) corrected) for eight main contrasts of interest with overlaid cortex parcellation from the human connectome project (HCP-MMP1.0). 66

Figure 21. Left posterior-lateral view of results of the univariate group level GLMX-analysis (with $p < 0.05$, Threshold free cluster enhancement (TFCE) corrected) for eight main contrasts of interest with overlaid cortex parcellation from the human connectome project (HCP-MMP1.0). 67

Figure 22. Number of contrasts showing significantly activated regions from univariate group level GLMX-analysis (with $p < 0.05$, Threshold free cluster enhancement (TFCE) corrected) from eight main contrasts of interest with overlaid cortex parcellation from the human connectome project (HCP-MMP1.0). 69

Figure 23. MNI-space location of single subject peak-t-scores for each contrast of interest in occipito-temporal cortex. Transparent ellipses depict 95%-confidence interval around the average location for the respective contrast A) 3D coordinates seen from horizontal plane, B) 3D coordinates seen from sagittal plane. 72

Figure 24. Left hemisphere regions of interest in LOTC area defined after the HCP-MMP1.0 cortex parcellation by Glasser et al. (2016) 74

Figure 26. Right hemisphere regions of interest in LOTC area defined after HCP-MMP1.0 cortex parcellation by Glasser et al. 2016. 76

Figure 27. Representational Similarity Matrix (RSM) construction, here shown for region MT. The same procedure was used in each participant and region of interest. First t-values are extracted from the ROI voxelwise for odd and even runs and resulting patterns of t-values (every tile corresponds to a voxel) are correlated across runs and contrasts resulting in one RSM (every tile corresponds to a correlation r of response patterns). 88

Figure 28. Average RSM construction. In every participant, t-scores were extracted from LOTC-ROIs for each run and each contrast of interest, odd and even runs were then correlated for participants individually within and across contrasts resulting in a representational similarity matrix (RSM). RSMs from each ROI were then averaged across participants. 89

Figure 29. Schematic depiction of model-RSMs for patterns reflecting a distinction of stimulus presentation A) in the left or right hemifield, B) in the upper or lower half of the visual field or C) for distinction of each quadrant. 92

Figure 30. Sensitivity for contrasts in region depicted by correlation of odd and even runs for each contrast of interest across ROIs. 94

Figure 31. Discrimination index (within contrast correlation- between contrast correlation) for each contrast of interest across ROIs. n.s.= not significant difference between within and between correlations.....	95
Figure 32. Average representational similarity matrices across participants in each region of interest in the left hemisphere. Single cells represent average correlation of distributed response patterns of experimental contrasts from odd and even runs depicting sensitivity and similarity of representations of domains in ROIs.	97
Figure 34. Dendrogram visualizations of the results from the left hemisphere based on hierarchical clustering of the correlations of distributed response patterns from experimental contrasts. Green box indicates recurring meaningful cluster of body related contrasts.	99
Figure 35. Average representational similarity matrices across participants in each region of interest in the right hemisphere. Single cells represent average correlation of distributed response patterns of experimental contrasts from odd and even runs depicting sensitivity and similarity of representations of domains in ROIs.	100
Figure 36. Difference between within contrast correlations and between contrast correlations for all regions of interest in the right hemisphere. *= $p < 0.001$; n.s.= not significant.	101
Figure 37. Dendrogram visualizations of the results from the right hemisphere based on hierarchical clustering of the correlations of distributed response patterns from experimental contrasts. Green box indicates recurring meaningful cluster of body related contrasts.	103
Figure 37. Similarity of RSMs across LOTC-regions for the domain contrasts. A) Schematic depiction of data used for analysis in red colour. B) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs. C) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs.	106
Figure 38. Average representational similarity matrices across participants in each region of interest in the right hemisphere for the spatial information experiment. Single cells represent average correlation of distributed response patterns of experimental conditions depicting sensitivity and similarity of representations for different locations of stimulus presentation.....	108

Figure 39. Similarity of RSMs across LOTC-regions for the spatial location contrasts. A) Schematic depiction of data used for analysis in red colour. B) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs. C) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs. 109

Figure 40. Model evaluation for response patterns to the spatial location experiment in LOTC-ROIs visualized on surface parcellation in the left hemisphere and bar chart for both hemispheres. A) Correlation of response pattern and hemifield-model. B) Correlation of response pattern and model of upper and lower visual field distinction. C) Correlation of response pattern and quadrant-model. 111

Figure 41. Left hemisphere whole brain functional connectivity maps from left hemisphere seed-ROIs in LOTC ($p < 0.001$, Bonferroni corrected, cluster size threshold of 40). 121

Figure 42. Right hemisphere whole brain functional connectivity maps from right hemisphere seed-ROIs in LOTC ($p < 0.001$, Bonferroni corrected, cluster size threshold of 40). 122

Figure 43. Similarity of functional connectivity (FC) maps across LOTC regions from left hemisphere seeds and right hemisphere seeds. A) Representational similarity matrix B) Dendrogram visualisation of results with hierarchical clustering. 124

List of Tables

Table 1. Regions involved in networks related to LOTC-domains 42

Table 2. List of Experiments, conditions, tasks, contrasts of interest for each experiment used during two sessions of MRI-scanning, the total duration of all runs together and number of runs. 47

Table 3. Contrasts overlapping in voxels with 5, 6 or 7 significantly active contrasts. 70

Table 5. Results for one-sample t-test for effects (uncorrected for multiple comparisons) of all contrasts in the regions of interest in the right hemisphere. T-values were extracted for each individual participant and region of interest (ROI) and averaged across participants for a mean t-value per contrast and region. 77

Table 6. Results for one-sample t-test to evaluate discrimination index (uncorrected for multiple comparisons) of all contrasts in the regions of interest in the left hemisphere.

Correlation on- and off-diagonal values were extracted for each region of interest (ROI) from the representational similarity matrixes..... 98

Table 7. Results for one-sample t-test to evaluate discrimination index (uncorrected for multiple comparisons) of all contrasts in the regions of interest in the right hemisphere. Correlation on- and off-diagonal values were extracted for each region of interest (ROI) from the representational similarity matrixes..... 102

1 Introduction

1.1 A hub for action perception in the brain

Daily social interactions of people highly depend on action, and understanding and responding to others' actions. An important function of the human brain facilitating these reactions and interactions is the perception of action. Action recognition can be achieved by combining diverse kinds of sensory input, from the shape of an effector like a body part or a tool, the motion involved, the spatial relation of effector and acted on object or person etc. and a possible change of all these over time, perspective, lightning and other external factors of actions. How these various inputs are processed in the brain, and especially how the information is combined in the human cortex, is understood only partly and can extend the understanding of more abstract action perception. For example, observing an action commonly involves rich and detailed information: the action of opening could be comprised of a water bottle (an object that is opened), a hand (a body part performing the action) and a movement (rotating the cap), but it might as well be comprised of a bin and a foot pushing a pedal to open it. The information therefore must be generalized across different features like a movement or body part of an action to understand the action on an abstract level.

One promising region for the association and integration of several of these components of action perception into a higher order representation of action is the lateral occipitotemporal cortex (LOT) (Lingnau & Downing, 2015). The LOT is the junction between the temporal, occipital and parietal lobes. Not only do different lobes of the brain assemble in this area, but also activation to a diverse range of possible components for action perception were reported for this region. Representations of motion (Papeo & Lingnau, 2015; Tootell, 1995), biological motion (Papeo & Lingnau, 2015; Lingnau & Petris, 2012), action observation (Grezes & Decety, 2001; Caspers, Zilles, Laird, Eickhoff, 2010), action performance (Astafiev, Stanley, Shulman, &

Corbetta, 2004; Orlov, Makin & Zohary, 2010), bodies (Downing, Jiang, Shuman, & Kanwisher, 2001), body parts (Orlov et al., 2010), tools (Chao, Haxby, & Martin, 1999) and verbs (Hafri, Trueswell & Epstein, 2017; Bedny, Caramazza, Grossman, Pascual-Leone, & Saxe, 2008) can be found in the area between superior temporal sulcus, lateral occipital sulcus, inferior temporal gyrus and halfway through the middle temporal gyrus (Lingnau & Downing, 2015). The detailed interrelation of these different domains in the LOTC area - specifically how they are spatially distributed and represented to facilitate information processing has not yet been explored.

This project aims to investigate the neural processes of these different domains to further the understanding of the cognitive architecture enabling the deduction of abstract action concepts. The first chapter summarizes the relevant domains that activate the LOTC and describes different principles of cortical architecture allowing for neuronal information processing. On that account, the location and function of the LOTC will be introduced in more detail in the following sections followed by an overview of possible organizational principles in the cortex.

The second chapter introduces the conducted MRI study in more detail summarizing general methods for functional MRI including eight different tasks, resting state and diffusion imaging. The third, fourth and fifth chapter focus on the topography and selectivity, representational content and connectivity within the LOTC that can be deduced from analysing our rich MRI dataset.

1.2 Definition of the LOTC

Lateral occipitotemporal cortex has been broadly defined by macroanatomy using mostly anatomical landmarks, in particular gyri and sulci. The LOTC can be seen as the area between

the lateral occipital sulcus (LOS), superior temporal sulcus (STS), the posterior half of the middle temporal gyrus (pMTG) and the inferior temporal gyrus (ITG) (Fig.1) (Lingnau & Downing, 2011).

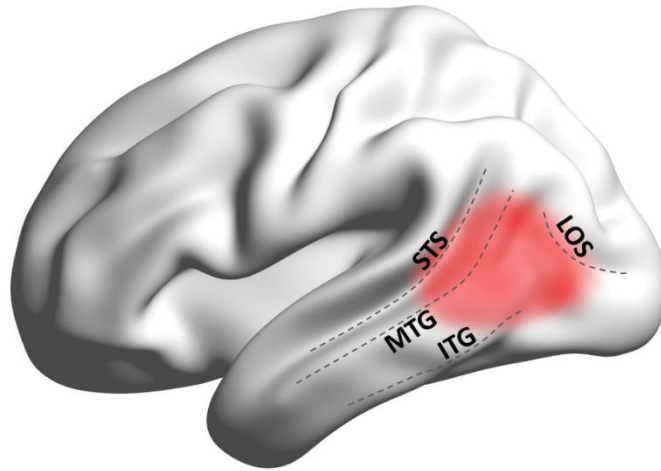


Figure 1. Approximate structural outline of the LOTC; LOS=Lateral occipital sulcus; STS=Superior temporal sulcus; MTG=middle temporal gyrus; ITG=Inferior temporal gyrus.

This structural definition can be used to circumscribe the extensive region, but the functional organization might not align with these structures and boundaries between neighbouring areas or subregions exist and can differ from the structural outline. Examining the consistency of the functionally defined area and structurally defined area would be beneficial to ensure a meaningful and sensible definition of this brain region in the future. Additionally, this definition is rather broad, and nomenclatures can differ between labs, using different atlases, or naming regions at different spatial scales, making a more precise exploration valuable to increase consensus between researchers and ease the transfer of information about the function and segmentation of this region.

Another way to define regions in the brain instead of cortical structure is by using functional localizer tasks. This method is based on the assumption that a region differs from surrounding areas based on the activity elicited within it by a task. One example is the extrastriate body area

(EBA) situated within the LOTC first described by Downing et al. (2001). To define EBA, a contrast of activity evoked by images of bodies and images of chairs can be used, highlighting increased activity within EBA for bodies compared to chairs. Functionally defined areas can be specified for each person by using the same task, but not exactly the same area will show increased activity: size and location can vary between participants (Tavor et al., 2016; Fox, Iaria & Barton, 2009), and sometimes they do not yield a significant response to define the region at all (Dumoulin et al., 2000). This definition has shortcomings as well because the functionally defined regions might depend on context (Friston, Rothstein, Geng, Sterzer & Henson, 2006) – for example EBA shows increased activity for bodies, but also for single body parts if the context is a task involving execution or observation of actions with these limbs (Orlov et al., 2010; Orlov, Porat, Makin & Zohari, 2014). If the same region is relevant for several different computations, possibly only a part of a bigger region is shown with a localizer. According to the results reported in chapter 2, a more complex structure like a topography is the underlying organizational principle with a gradual change instead of sharp boundaries.

1.3 Activation of LOTC - Overview of diverse tasks activating LOTC

To better understand the function of the LOTC, an initial important step is to investigate what stimuli or tasks are represented in the area. To explore the spatial distribution and convergence of different tasks in the LOTC, different localizers from diverse research areas were used, that all activate a similar area -namely LOTC. The next section briefly summarizes the background of each concept underlying mostly established localizer tasks. The shown findings include a wide range of features of actions, from lower-level features like motion to more complex features like body parts, but even though the categorization of a hand as such might be more complex as noticing motion, both are key features to categorize an action as such. As these features differ

in their complexity the tasks to ensure the person engages in for example motion detection, body part categorization or action execution cannot be the same. Yet to investigate the interplay of features and concepts on different levels of complexity, we focus on the possible hub region LOTC. Figures 2-9 show the results of meta-analyses for respective terms done in neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>). This online platform contains data from thousands of published MRI papers and allows for an automated meta-analysis of this body of research. A term or topic can be chosen and all relevant results from papers related to the chosen keyword are then used for a meta-analysis, outlining brain regions related to the term.

1.3.1 Basic motion perception

A motion sensitive area in the temporo-parietal-occipital junction was first described in macaque monkeys (Maunsell & Van Essen, 1983) showing tuning of single neurons to the direction and speed of a stimulus. Zeki et al. (1991) was able to also show this in the human brain, using positron emission tomography (PET) following single cell recordings in area V5 (also called middle temporal complex MT). To evaluate the motion sensitivity, Zeki et al. (1991) showed stimuli consisting of a large number of black squares on white background either moving in one of eight directions or static.

By contrasting the blood flow measured during the moving condition with the blood flow during the static condition, they were able to show that human V1/V2 and V5 responded stronger during the presentation of moving in comparison to static stimuli. A few years later the functional specialization of MT was also demonstrated with fMRI (Tootell et al., 1995) using moving (expanding or contracting) white dots on black background compared to stationary dots. Today it is generally accepted that MT – also referred to as hMT+/human V5- is involved in the processing of motion and it's direction or speed in the visual receptive field (Born & Bradley, 2005; Hock & Nichols, 2013; Lingnau, Ashida, Wall, & Smith, 2009). The hMT+ can be defined in

individual participants by using a basic motion localizer task, comparing the response to dots that move on the radial axis with static dots (Tootell et al., 1995; Lingnau et al., 2009; Papeo & Lingnau, 2015).

Even though the role of hMT+ in motion perception is well documented, its function seems to go beyond mere perception and involves integration and segmentation of temporally separated inputs and therefore reduction of the complexity of different kinds of incoming information (Born & Bradley, 2005). Motion can only be perceived over time, as the change of location from one time-point to the other is key to induce motion perception. As the reviewed research has shown, area MT/hMT+ is highly relevant for the perception of motion, this is also supported by a meta-analysis using neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>) showing bilateral clusters in the LOTC area for the term motion from 451 studies (see Figure 2. for left hemisphere cluster).

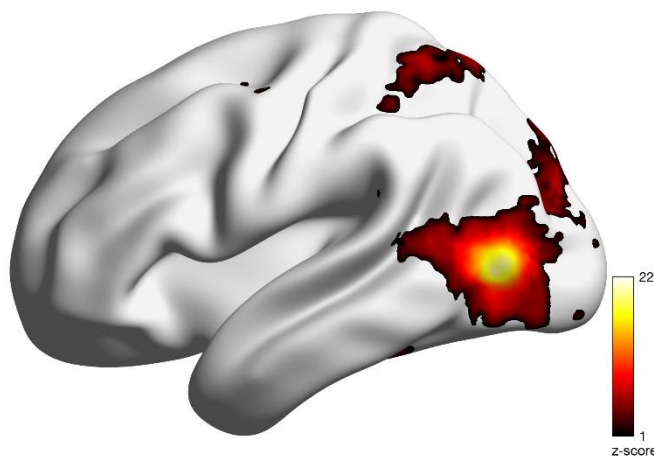


Figure 2. z-map for neurosynth meta-analysis of the term 'motion'

1.3.2 *Biological motion perception*

More lateral and anterior than hMT+, a region at the ventral-occipital section of the superior temporal sulcus (STS) has been linked to the visual perception of whole-body motion, often referred to as biological motion (Grossman et al., 2000).

To study the coherent motion of humans, the impression of a moving body can be induced with a point-light display of the movement of main joints of a person (Johansson, 1973), reducing the complexity of the stimulus to essential points. To localize areas sensitive to biological motion, the activation initiated by point-light displays can be compared to the response to scrambled animations of points (Berenthal & Pinto, 1994; Neri, Morrone & Burr, 1998; Grossman et al., 2000; Lingnau & Petris, 2012; Papeo & Lingnau, 2015). There is evidence for a right lateralization of biological motion processing (Grossman et al., 2000; Peuskens et al., 2005). While Peelen et al. (2006) did not find meaningful representations of biological motion in areas sensitive to basic motion, the representations of bodies and point-light displays in areas sensitive to human bodies showed high correlation, indicating a similar underlying neuronal response for static bodies and moving point-light displays.

A meta-analysis on human movement perception indicated activation in the LOTC and fusiform cortex – containing also previously described area hMT+ – for biological motion (Grosbras, Beaton & Eickhoff, 2012). However, it is unclear how to interpret the overlapping for basic motion and biological motion as revealed by Grosbras et al. (2012). As an example, Peuskens, Vanrie, Verfaillie & Orban (2005) did not find overlapping regions when comparing basic motion and biological motion directly. Likewise, Peelen, Wiggett & Downing (2006) showed that the hMT+, even though activated by biological motion, has no functional role in its perception.

1.3.3 *Body perception*

More ventrally than motion sensitive regions, an area sensitive to the observation of the human body, the extrastriate body area (EBA), is situated in the posterior inferior temporal sulcus and middle temporal gyrus. Various kinds of depiction of bodies like static images, stick figures or silhouettes are linked to higher activity in the EBA in comparison to images of objects, tools and faces (Downing et al., 2001). The EBA also showed a preference for intact whole-body configurations compared to the sum of body parts indicating that the configuration of bodies plays a role in body perception (Brandman & Yovel, 2014).

To localize the EBA, participants are commonly shown pictures of human bodies (without heads) and chairs to contrast the activation. To visualize the approximate location of a body-sensitive area, a meta-analysis for the term body using neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>) is shown in Figure 3, revealing a strong cluster in the LOTC.

EBA has also been shown to be sensitive to biological motion stimuli but only processing the body form and not the motion (Downing, Peelen, Wiggett, & Tew, 2006). The active movements of limbs to perform an action modulates EBA as well (Astafiev et al., 2004). Moreover, the area seems to be involved in representing both seen bodies of other people and action execution with the own body even without visual feedback (Astafiev et al., 2004). An overlap of motion sensitive hMT+/V5 and body sensitive EBA was found in retinotopically defined MT regions, where motion and shape cues could be combined to facilitate abstraction of action concepts (Ferri, Kolster, Jastoff & Orban, 2013).

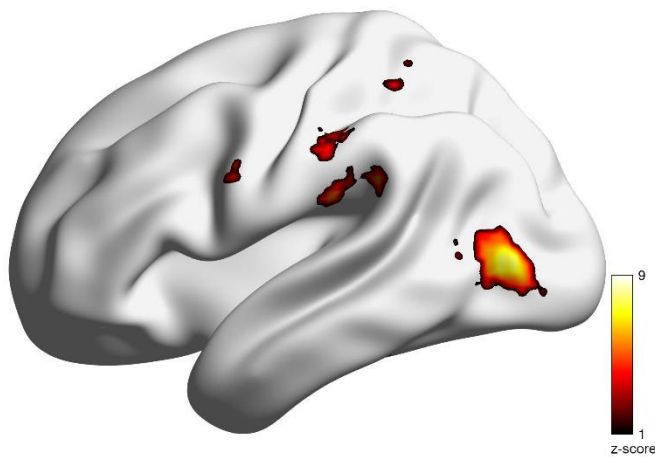


Figure 3. z-map for neurosynth meta-analysis of the term 'body'

1.3.4 *Body parts*

Downing et al. (2001) also found that, like whole bodies, body parts activate the LOTC significantly more than object parts, both for photographs and drawings. Investigating the specialization of the LOTC for visually presented body parts more precisely, Orlov et al. (2010), showed topographically organized clusters selective for different body parts like upper or lower limbs, trunks and faces. Bracci, Ietswaart, Peelen and Cavina-Pratesi (2010) also reported an area in extrastriate cortex preferring hand stimuli over other body parts or whole bodies.

The underlying principle of the organization of the representations of body parts in the occipito-temporal cortex could be the functional-semantic properties of them. Body parts like hands and feet are key in action execution, faces convey social cues, while parts like the chest and waist are less meaningful. To test this, Bracci, Caramazza and Peelen (2015) compared the neural representation of body parts with different features they might be organized by – physical shape, perceived shape, physical proximity, cortical homunculus or semantic similarity. The best model to explain the organization of brain activity was based on the semantic similarity resulting in a tripartite organization of body part selective areas into clusters of action effectors, face parts and non-effector body parts (Bracci, Caramazza, & Peelen, 2015). A meta-analysis on faces,

hands and feet using neurosynth revealed the maps shown in figure 4, indicating activations for faces and hands but not for feet in the LOTC area.

The spatial relationship between body part sensitive areas and motion sensitive areas was the subject of several studies where an extensive overlap was found (Downing, Wigget & Peelen, 2007; Peelen & Downing, 2007). However, a more recent study using high resolution fMRI indicates distinct areas for motion and a crescent-shaped tripartite topography for body parts surrounding hMT+ and MST (Weiner & Grill-Spector, 2011).

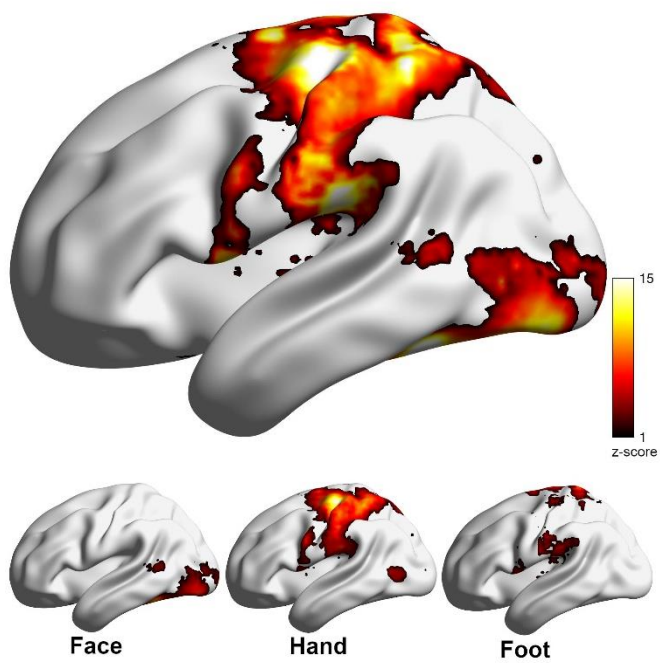


Figure 4. z-map for a combination of neurosynth meta-analyses of the terms 'face', 'hand', 'foot' and each term separately.

1.3.5 Tools

Looking at pictures of tools, naming or reading tool names, and the perception of tool motion, activates the posterior part of the middle temporal gyrus (Chao et al., 1999; Beauchamp, Lee, Haxby, & Martin, 2002) (see Fig. 5 for a neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>) meta-analysis on the term tool). Neural representations of tools and hands were found to overlap in the left LOTC cortex possibly, reflecting the shared object manipulation properties (Bracci, Cavina-Pratesi, Ietswaart, Caramazza & Peelen, 2012). Furthermore, even though the tool sensitive area is located anterior to hMT+ (Bracci et al., 2012; Beauchamp, 2002), it shows increased activation during visually presented motion that is typical for hand tool actions (Beauchamp, Lee, Haxby, & Martin, 2003; Beauchamp et al., 2002).

In contrast to the idea of shared representations of hands and tools, using multivariate pattern analysis, Gallivan, McLean, Valyear and Culham (2013) found that the execution of hand actions was coded in hand and body areas while tool actions could be decoded using signals from tool selective areas. The only areas where Gallivan et al. (2013) found a shared representation of hands and tools were frontoparietal areas, whereas representations in the LOTC were found to be more selective and less integrated. These results differ but as Bracci et al. (2012) used visual stimuli while Gallivan et al. (2013) had participants execute an action, this divergence might indicate separate representations for visual processing of hands and tools and responses for the execution of hand and tool actions. Looking at three separate tool and hand sensitive areas, yet another picture emerged indicating different representations at different locations: tool action-related representations were found in left IPS, category specific representations in ventral occipitotemporal cortex (VOTC) and left LOTC- possibly acting as a hub region- encodes both action and category information (Bracci, Cavina-Pratesi, Connolly, & Ietswaart, 2016). Therefore, the LOTC is an important region for tool processing, possibly combining the information of belonging to a category like tool or hand with a respective action concept.

The sensitivity of tool areas does not seem to depend on visual stimulation though. As summarised in a review from Lewis (2006), studies reported tool-related activity in left LOTC for hearing tool names or sounds, reading of tool words and pantomiming or imagery of tool use. Tool sensitivity in LOTC was also found in congenitally blind participants auditorily presented with words referring to tools, signifying the development of the sensitivity is independent from visual experience (Peelen, Bracci, Lu, He & Caramazza; 2013).

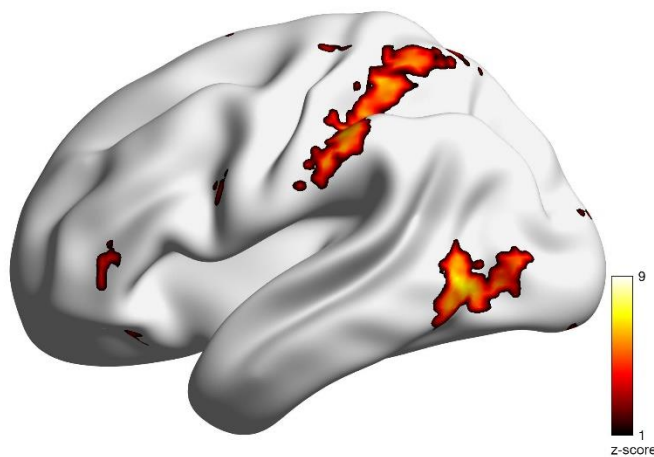


Figure 5. z-map for neurosynth meta-analysis of the term 'tools'

1.3.6 Action observation

Areas in the LOTC do not only encode category information like tools and hands but were also found to represent more abstract information about how and whether the object can be manipulated and used for an action, so both information about what category (animate/inanimate) an object belongs to and how much action related information the object

has (nonmanipulable/tool) are represented in LOTC (Bracci et al., 2016). Tasks that require that participants perceive and/or understand actions might engage the LOTC.

Early research on action observation in the human brain was conducted with Positron Emission Tomography. Rizzolatti, Fadiga, Gallese & Fogassi (1996) compared the observation of a grasping action with static objects and found activation in the left middle occipital and temporal gyrus. The involvement of bilateral lateral occipital areas and pMTG in action observation was also found in two meta-analyses (Grezes & Decety, 2001; Caspers et al., 2010). Figure 6. depicts the resulting z-map from a meta-analysis in neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>) on the term action observation. The occipitotemporal cortex was found to represent observed actions at generalized and therefore abstract levels instead of concrete exemplars of action (Wurm & Lingnau, 2015) meaning it does not differentiate between details like the object acted upon, but rather the more abstract concept of, for example, opening vs. closing is coded in OTC. The posterior temporal cortex was also found to represent information about actions instead of specific body parts, movements or objects relevant for the action (Vannuscorps, Wurm, Striem-Amit & Caramazza, 2018; Wurm & Lingnau, 2015) and generalizes across modes of stimulation like photographs, videos and written descriptions (Hafri et al., 2017).

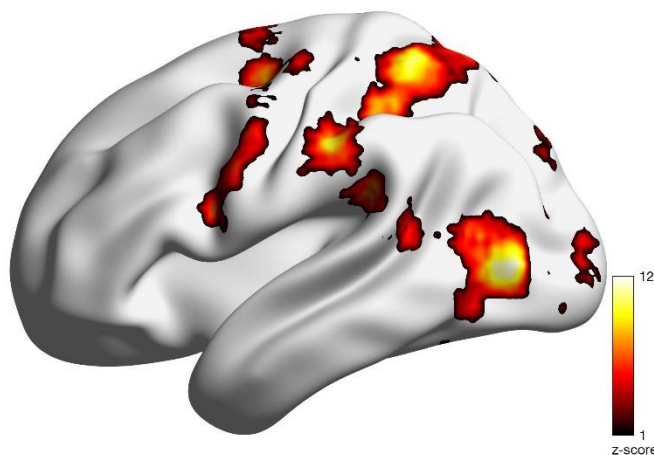


Figure 6. z-map for neurosynth meta-analysis of the term 'action observation'

1.3.7 Planning & Executing actions

Astafiev et al. (2004) asked their participants to execute hand or foot actions within the fMRI scanner. They obtained activations for limb movements in the EBA that could not be explained by visual perception of body parts, showing that not only motor areas are involved in the planning and execution of actions. In a study mapping different body parts, Orlov et al. (2010) found overlapping activation of visually presented body parts and the unseen movement of body parts in occipitotemporal cortex. Likewise, Kühn, Keizer, Rombouts & Hommel (2011) found activation for preparation of a manual action in the EBA. To rule out the possibility that action related activation in occipitotemporal areas were only due to tactile or visual feedback, Gallivan et al. (2013) investigated a preview, planning and execution phase of hand movements. From lateral occipital regions they were able to decode what action (grasp or reach) with what hand (left or right) was about to be executed so the occipitotemporal region was involved already in the planning of a movement and responded independently from any visual or tactile information from the movement. Also looking at the planning of actions, the decoding of action plans in the right pMTG was found to be only possible if a choice between actions was given, meaning the LOTC might be involved in the selection of actions (Ariani, Wurm, & Lingnau, 2015).

Action performance or execution were not available on neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>), but a meta-analysis for the term *movement* shows a cluster in the LOTC area in figure 7.

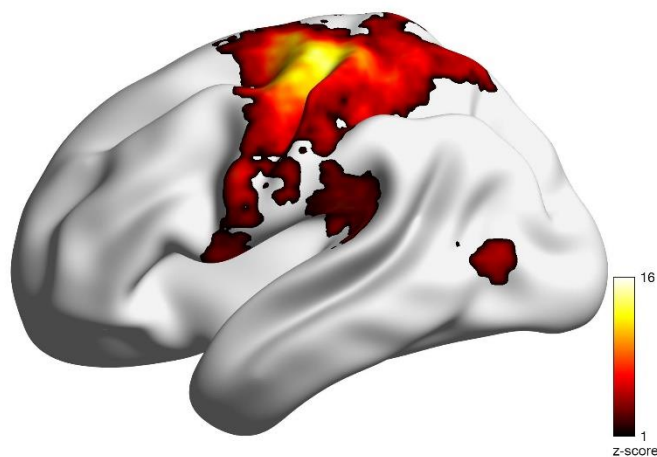


Figure 7. z-map for neurosynth meta-analysis of the term 'movement'

1.3.8 Verbs

Action verbs are of interest for the function of LOTC as well, as they commonly refer to actions. Responses to verbs were found in the right posterior-lateral-temporal cortex (pLTC), close to LOTC (Bedny et al. 2008; Kable, Lease-Spellmeyer & Chatterjee, 2002). The contrast of verbs > nouns reliably recruits the left PLTC (Bedny et al. 2008; Kable, Lease-Spellmeyer & Chatterjee, 2002) but it is still discussed what properties of verbs are the basis of this difference. One branch of research found that the property of binding sentence components into a unified structure to create a meaningful statement (predication properties) of the grammatical class of verbs might drive the activation and not the semantics (event vs. state) of words (Hernandez, Fairhall, Lenci, Baroni & Caramazza, 2014; Peelen, Romagno & Caramazza, 2012). Additionally, studies found an effect of semantics where event related verbs but also event nouns showed larger responses than object nouns in left pMTG indicating an event related representation (Bedny, Dravida & Saxe, 2014).

Overlapping activation for verbs were reported in hMT+ related to basic motion (Saygin, McCullough, Alac, & Emmorey, 2010), and in pSTS and LOTC associated with biological motion (Bedny et al., 2008; Deen & McCarthy, 2010). The three overlapping regions might encompass

two different representations – while MT and pSTS show higher activity for verbs compared to nouns but no modulation of this effect by semantic content like implied motion (like PLTC does), there could be two dissociable representations: one of grammatical class in MT and pSTS and one of events/actions within other regions of the LOTC area (Papeo & Lingnau, 2014, Peelen, Romagno & Caramazza, 2012). Action verbs related to specific body parts like pick (hand) or kick (foot) were found to activate regions in motor cortex and premotor cortex in a somatotopic map similar to the somatotopic map of performed actions (Hauk, Johnsrude and Pulvermüller, 2004) supporting a representation of words in a distributed map or topography according to their meaning. In line with the reported findings, a meta-analysis for the term verbs results in a map including a cluster in posterior temporal cortex (see Figure 8 for neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>) meta-analysis for *verbs*).

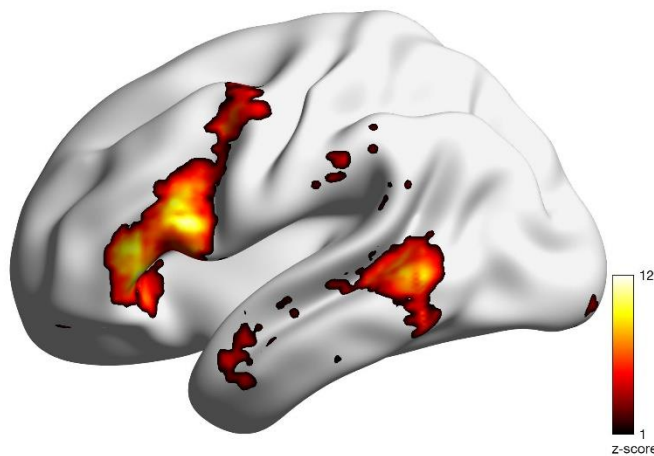


Figure 8. z-map for neurosynth meta-analysis of the term 'verbs'

1.4 The hub in lateral occipitotemporal cortex

Understanding actions is a complex process and a range of different kinds of information can contribute to the interpretation of a combination of features as one specific action. Motion, involvement of body parts and tools can be characteristic for an action but is the sum of the features represented in a distinct area than the single features? For example, is a grasping action represented by activating the hand area and the area recruited during the processing of basic motion? All together the LOTC is associated with the computation of the domains introduced above, i.e., basic motion, tools, body parts (face and hand – not foot), bodies, action observation and execution. Using neurosynth (Yarkoni, 2014, <http://old.neurosynth.org/>), all these terms show overlapping clusters in the LOTC area (Fig. 9). This overlapping cluster did not overlap with the clusters generated for the term 'verbs', as the cluster for verbs was located more anteriorly in temporal cortex, whereas the term 'biological motion' was not available in neurosynth. This meta-analysis already indicates that for example the processing of hands and basic motion to some extent recruits the same area. However, to better understand the spatial and functional relation of action-related information, this needs to be investigated in more detail as several possible organizing principles could apply to representations in the LOTC.

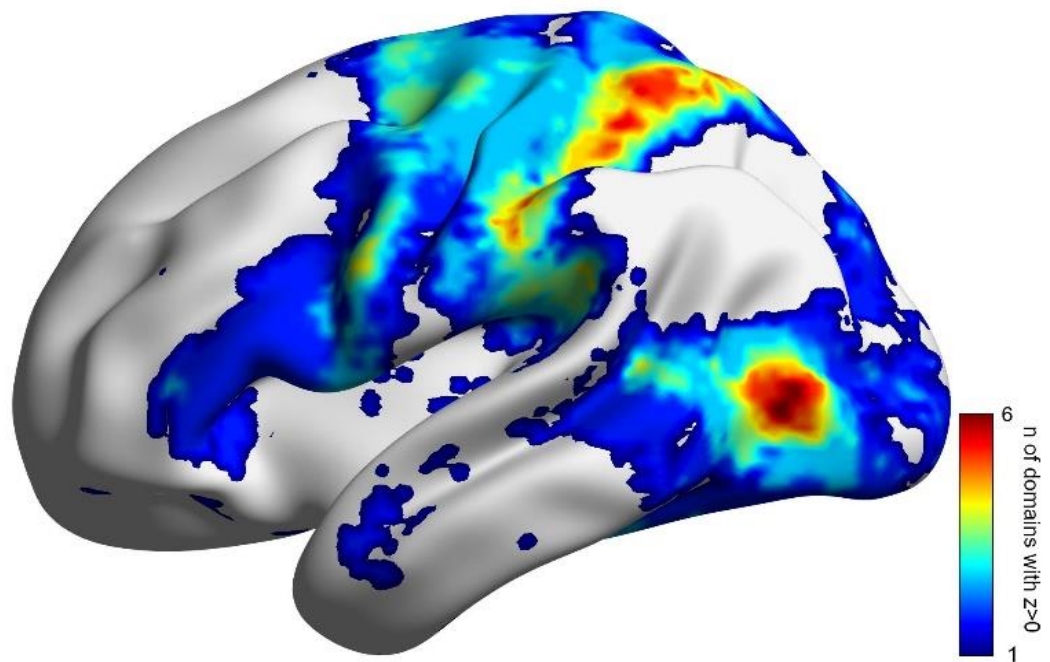


Figure 9. Combination of binarized z-maps from neurosynth meta-analysis of the terms 'motion', 'body', 'face', 'hand', 'foot', 'tool', 'action observation', 'movement', 'verb' showing overlap of the maps.

1.5 Organizational principles

The LOTC is involved in the processing of various aforementioned inputs and the question arises how the representations are organized functionally and structurally. In this section, an overview of proposed organizing principles of the cortex will be provided.

The architecture of the cerebral cortex is well studied in some parts of the brain – especially sensory areas like visual and somatosensory cortices. In primary visual cortex for example the visual input is mapped onto the cortex directly in the so called retinotopic map. The array of the input on the retina directly corresponds to the array of representations in early visual cortex. This means that an object in the center of the visual field is represented in the center of visual

cortex while the periphery of the visual field corresponds to visual regions further away from the center resulting in an orderly eccentricity map. This is combined with a map of the polar angle representing the position of the object in the visual field such that objects presented in the upper left quadrant of the visual field are represented in the ventral portion of the right visual cortex. Furthermore, objects right next to each other in the visual field would also lead to responses of visual areas next to each other. Likewise, the somatosensory cortex shows the correspondence of the position of the receptive field on the surface of the skin from different body parts and the position of neurons representing these fields are called somatotopic map. However cortical organization is less clear in higher order association areas like the LOTC, where a function could be to combine different information into more complex representations and pass on information between regions (Lingnau & Downing, 2015), as rich and diverse information seems to converge in this area.

1.5.1 Modules

One of the most primary forms of cortical organization that has been proposed is the subdivision of the cortex into discrete regions dedicated to specific functions, termed modules (Fodor, 1983; Fodor, 2001; Caramazza & Shelton, 1998; Kanwisher, 2010). Neurons that have similar properties tend to cluster together in the brain, so the concept of modules often refers to a region of cortex, with a higher density of connections between neurons within the module than to other neurons, modules or networks (Sporns & Betzel, 2016). Modules are assumed to be functionally specialized to selectively compute one type of information or fulfill one single purpose and can be seen as a processing unit, for example for the processing of bodies in the extrastriate body area (EBA)(Downing et al., 2001), for faces in the Fusiform Face Area (Kanwisher, McDermott & Chun, 1997; McCarthy, Puce, Gore & Allison, 1997), for places in the

Parahippocampal Place Area (Epstein & Kanwisher, 1998) or motion sensitive area hMT+ (Tootell et al., 1995). In neuroscience, two competing views are much debated directly relating to modules and the underlying organization of the mind: modules are either domain-specific or domain-general (Fodor, 1983, Kanwisher, 2000). According to domain specificity, modules are specific to the information that is processed by them, for example there is a separate module categorizing bodies and one detecting motion. The domain – or category of the input determines, in what module it is processed. Opposing this view, the domain-general module would be specific to a function or process that needs to be carried out like categorization or volition. The kind of process that is applied to the input determines in what module it is processed (Fodor, 1983, Kanwisher, 2000).

1.5.2 Hierarchy

Going back to the early visual cortex, a hierarchical series of modules was suggested, each processing visual input at an increasing level of complexity (Marr, 1982; Felleman & Van Essen, 1991). The visual cortex encodes sensory information in a hierarchical manner from simple contrast analyses to determine edges and lines, to an assembly of edges creating shapes, and later objects with the information being passed on hierarchically from posterior occipital cortex to more anterior regions. In the striate cortex for example complex cells of lower order are sensitive to line stimuli of a specific orientation, while more anterior higher order hypercomplex cells are known to receive input from several of these cells about the line orientation thus decreasing specificity but increasing the receptive field (Hubel & Wiesel, 1965). In a hierarchical arrangement like this, regions are selective for properties of input with a gradient of, for example, complexity (Petersen & Sporns, 2015).

A hierarchical information transmission would lead to the prediction of strong connections between neighboring areas, while long range connections would be expected to be less pronounced (deHaan & Cowey, 2011; Kravitz Saleem, Baker, Ungerleider & Mishkin, 2013). Object recognition as well as action recognition can be achieved by a hierarchy of representations from basic levels of visual perception to categorization of objects or actions (Riesenhuber & Poggio, 1999; Grafton & Hamilton, 2007) implying a gradual correspondence between level of complexity and level of cortical processing of the module.

1.5.3 Topography

Similar to a hierarchical mapping but gradually mapping input for just one feature onto the cortex, topographic maps were established in the cortex. The first topographic map found in the human brain is the so-called homunculus in primary motor and somatosensory cortex (Ferrier, 1874; Penfield & Boldrey, 1937; Kwan, MacKay, Murphy & Wong, 1978). The homunculus is a representation of the body in the brain, organized systematically so that cortical regions correspond to body parts (also called somatotopy). The early visual areas also show topographic maps, such as the retinotopic map of retinal input with a systematic relationship between angular position and eccentricity of visual information and neuronal representation of this spatial information (Hubel & Wiesel, 1962) where the array of receptors on the retina relates to the array of the representation on the cortex.

Borders of visual areas V1-V4 can be defined by retinotopic mapping indicating a gradual change of sensitivity for location on the retina across early visual cortex. Within the receptive visual field, checkerboard patterned stimuli are moved, either rotating a wedge or dilating a ring to activate different angular or eccentric positions. Neural representations in early visual areas are strongly related to input from the retina (Gazzaniga, Ivry, & Mangun, 2014). Standard

procedures for retinotopy are based on early experiments (Sereno, McDonald, & Allman, 1994; Sereno, Dale, Reppas & Kwong, 1995) matching the temporal phase of the moving shapes with neural activation in early visual cortices of monkeys and humans.

In higher level cortex, Malach, Levy and Hasson (2002) reported a similar, less specific topography like an eccentricity map. They found a bias for the location of representations for objects related to the resolution needs for the identification of objects. Objects that are categorised based on fine details from high-resolution, foveal representations like faces or words are associated with a central-field bias of the representations, while objects like buildings need more large-scale integration of peripheral, low-magnification representations and are related to peripheral-field biased representations (Hasson, Levy, Behrmann, Hendler, & Malach, 2002; Malch, Levy & Hasson, 2002). As the information in higher cortical areas is less directly linked to the retinotopic perception but might follow not only a eccentricity bias but also a gradient, like Roth & Zohary (2015) proposed for the dorsal stream of visual processing. They found less information about the stimulus position to be encoded in higher, more anterior areas, whereas information about identity of tools or hands got enhanced indicating a gradual transition from direct mapping of information to a more abstract and categorical representation. A similar gradient could also be found in the ventral stream considering LOC in the ventral stream and pIPS in the dorsal stream both similarly represent objects independent from the viewpoint or size (Konen & Kastner, 2008).

The occipitotemporal cortex was also shown to be organized by a gradient of real-world size (Konkle & Oliva, 2012) with a medial-lateral organization of big-small sized objects independently from the size of the object on the retina or whether the objects were seen or imagined.

Topographic maps can not only represent two dimensions like it was shown in primary visual cortex with the mapping of the input of the retina (which is also 2D), but it can also decode more dimensions – even from different categories at the same time (Aflalo & Graziano, 2006). Graziano, Aflalo and Cook (2005) and Graziano, Taylor and Moore (2002) found a map representing the goal location of the hand – when stimulated at a specific area, a hand movement to a certain location in 3D space was executed by a monkey. A possible organizational principle was proposed, where multiple maps are superimposed or intermeshed to form one hybrid map (Graziano et al., 2005) for example in monkeys the superimposition of somatotopic information, function of an action and spatial location of relevant body parts.

Likewise, in LOTC, different objects or categories have been suggested to be represented by combined, overlapping maps (Op de Beeck, Haushofer & Kanwisher, 2008) like a shape map (Vernon, Gouws, Lawrence, Wade & Morland, 2016; Op de Beeck, Baker, DiCarlo & Kanwisher, 2006) overlapping with a map corresponding to the connections of areas to other brain regions (Mahon et al., 2007). The combination of these maps could lead to a strong selectivity where only stimuli of a particular combination like a round and elongated shape and connectivity to motor areas indicate a body – activating both maps in a specific way giving way to a body-selective activation overall (Op de Beeck et al., 2008).

A complex multidimensional topography could be a property of a module, but it could also encompass a wider area of cortex. An example would be retinotopy, where sensitivity to spatial location in the visual field is strong in early visual areas like V1 and gets gradually less specific (Roth & Zohary, 2015) but is still present in lateral occipital areas (LO1 and LO2) (Larsson & Heeger, 2006), LOC, EBA and MT (Golomb & Kanwisher, 2011). An additional proposed gradient from early visual cortex to more anterior ventral cortex is the transition from concrete low-level visual input to abstract high-level object representation (Grill-Spector & Malach, 2004; Roth & Zohary, 2015), generalizing across more basic visual features of an object to enable for example

categorization independent from viewpoint, colour, texture, or shape. While information about the location is gradually lost, information about object identity is enhanced (Roth & Zohary, 2015), yet information is not solely coded in one small, circumscribed module but rather spread across the cortex.

As other topographic gradients were established in the LOTC area like representing object size, animacy of shown objects (Konkle & Caramazza, 2013), a topography of body parts (Orlov et al., 2010) or gradients of how social, object-related and abstract observed actions are (Wurm, Caramazza & Lingnau, 2017) the LOTC might be a hub area where multiple features of action representations are coded by a representational space with multiplexed gradients describing key aspects of action in combination (Lingnau & Downing, 2015). Cortical hubs were proposed as regions receiving and integrating information to create more general concepts not restricted to one modality or input (Lambon, 2014; Bajada, Trujillo-Barreto, Parker, Cloutman & Ralph, 2019) and LOTC might be a key area to integrate multimodal information regarding actions into abstract concepts of action.

1.5.4 Distributed Processing

Another proposed principle is that information in the cortex is represented in a distributed way (Haxby et al., 2001). The idea is that information is captured by the pattern of activation in a representational space instead of one single modular area and this space can be defined by several features of the represented information (Haxby et al., 2011; Kriegeskorte, Mur & Bandettini, 2008). The representation of information might not be realized by just one constrained region but could as well be superimposed on a wider part of the cortex and might form a continuous map (Graziano et al., 2005; Op de Beeck et al., 2008).

This concept is also not only in line with the description of the function of one brain region but can also be extended to a non-contiguous representation realized by the interplay of several regions – a network. For example, the LOTC was found to be part of the action observation network (AON) in both humans and monkeys facilitating the understanding of seen actions (Nelissen et al., 2011; Grezes & Decety, 2001; Caspers et al., 2010). The AON includes the ventral part of premotor cortex or inferior frontal area (IFG), inferior parietal lobe (IPL) and posterior middle temporal gyrus (MTG) which is part of the wider LOTC and action representations might be encoded by the distributed network of neurons (McClelland & Rogers, 2003) like it has already been suggested for face perception (Freiwald & Tsao, 2010; Pitcher, Walsh, & Duchaine, 2011).

The LOTC or subareas of the LOTC might be organized by a topography or gradient relating to the connections to other networks and areas as proposed by the distributed domain-specific hypothesis (Mahon & Caramazza, 2011). In this account, the location of a domain-specific area is constrained by the connections it has, for example, the tool sensitive regions would be located in an area that is connected to regions involved in manipulation of objects (Mahon et al., 2007). The connectivity of LOTC, or subregions of it, might therefore be of interest to better understand the underlying function.

Areas V4 and TEO (posterior inferior temporal cortex) in primates situated in the anterior ventral part of LOTC for example show a topographic connectivity pattern where the location in the temporal lobe corresponds to the location of the target areas in the neostriatum – anterior areas in temporal cortex are connected to anterior parts of the neostriatum and both gradually changing over to more posterior areas in temporal cortex and neostriatum in the macaque brain (Kravitz et al., 2013).

The MT/MST region situated in the posterior dorsal part of LOTC on the other hand does not show a topographic connectivity pattern but it is heavily interconnected with the surrounding occipitoparietal circuit and also connects to prefrontal areas which supports spatial working memory (Kravitz, Saleem, Baker & Mishkin 2011). This indicates different subareas of LOTC related to the processing of different features also show different interconnections. In line with this idea, Turken & Dronkers (2011) investigated the functional connectivity of different parts of posterior MTG (overlapping with LOTC) and found a distinction between posterior parts of MTG showing mainly connectivity to closest surrounding areas while the anterior parts of MTG exhibited a more widely distributed connectivity. This could indicate different functions of anterior and posterior regions and while the information processed in posterior regions might be processed more locally for example in a hierarchical order, the information in anterior regions could be more independent and abstract from visual input and the hierarchy arising from visual cortex.

1.5.5 Networks

Not only action observation is related to a network of coactivated or correlated brain regions found during functional or resting state scans. Looking into other domains activating the LOTC, similar networks were reported for visual motion (Culham, He, Dukelow & Verstraten, 2001), biological motion (Grossman, Jardin & Pyles, 2010; Saygin, 2007; Saygin, Wilson, Hagler, Bates & Sereno, 2004), tools (Peelen et al., 2013; Reynaud, Lesourd, Navarro & Osiurak, 2016; Lewis, 2006; Watson & Buxbaum, 2015), bodies (Ramsey, 2018; Hodzic, Kaas, Muckli, Stirn & Singer, 2009; Amoruso, Couto & Ibanez, 2011), action observation (Caspers et al., 2010; Grezes & Decety, 2001; Hardwick, Caspers, Eickhoff & Swinnen, 2018), action performance (Hardwick et al., 2018, Gerardin et al. 2000; Orlov et al., 2010; Gallivan, McLean, Valyear, Pettypiece &

Culham, 2011) and language (Turkeltaub, Eden, Jones & Zeffiro, 2002; Mechelli et al., 2005; Fedorenko & Thompson-Schill, 2014; Hagoort, 2014) (See summary of networks in Tab. 1 and Fig. 10).

Evaluating the different networks, most of them include similar areas (Table 1) with inferior frontal gyrus, the premotor cortex, inferior and superior parts of the parietal lobe, superior temporal cortex, inferior temporal or fusiform cortex and all of them include regions in the LOTC. The tool and body (Ramsey, 2018) networks also extend into prefrontal areas, body and language networks also include temporal pole, the visual motion network (Culham et al., 2001) encompasses lingual gyrus and V1, while action performance extends into cerebellum and basal ganglia and includes M1 and somatosensory cortex.

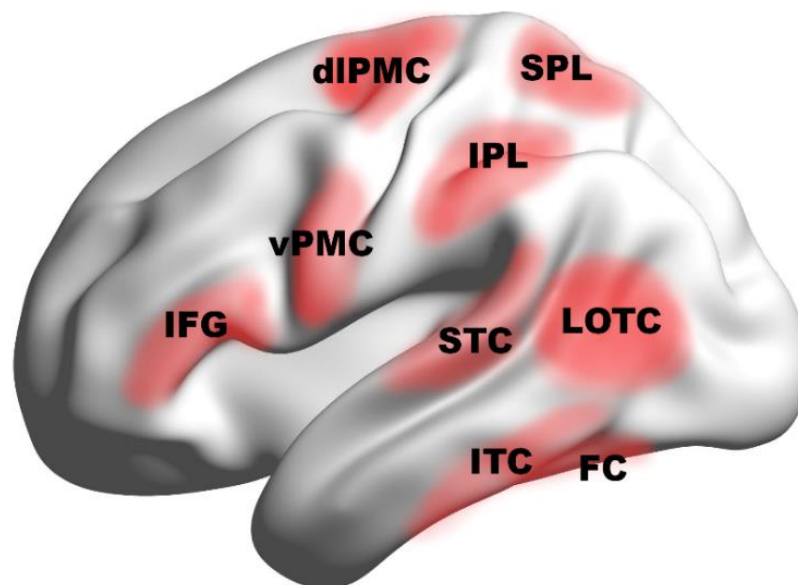


Figure 10. Location of main regions involved in networks related to domains activating the LOTC. IFG=Inferior frontal gyrus; vPMC = ventral premotor cortex; dIPMC= dorsolateral premotor cortex; SPL= superior parietal lobe; IPL= inferior parietal lobe; STC= superior temporal cortex; LOTC = lateral occipitotemporal cortex; ITC= inferior temporal cortex; FC = fusiform cortex

The reported domains all share one node in their networks- namely LOTC - but they are also highly overlapping in several regions while, as shown in Table. 1, they all have unique networks containing different combinations or different additional regions. Different subareas of the LOTC thus might be part of different networks to facilitate the processing of the diverse domains and the connectivity – functional and structural as Turken and Dronkers (2011) and Hutchison, Culham, Everling, Flanagan and Gallivan (2014) have shown for parts of MTG and OTC – can give information on the underlying organizational structure in the LOTC. Turken and Dronkers (2011) investigated the language comprehension network using functional and structural connectivity of MTG and subregions of MTG, and found a gradient with extensively distributed connectivity in more anterior MTG-subdivisions and less widely distributed patterns of connectivity in posterior subdivisions. Also using resting-state functional connectivity, Hutchison et al. (2014) used several localizers to identify seed-regions of interest based on selective responses to object-, face-, scene-, body- and tool-stimuli and even though the seed-regions lie close together and some respond to the same category (for example EBA and FBA), no consistent relationship between proximity of ROIs and similarity in connectivity pattern was reported and only little overlap of whole-brain connectivity pattern was found in their study for same category ROIs. Further research is needed to address these inconsistent findings and investigate the functional connectivity of LOTC.

Table 1. Regions involved in networks related to LOTC-domains

ROI	Motion ^a	Biological Motion ^b	Tool ^c	Body ^d	Action Observation ^e	Action Performance ^f	Language ^g
PFC			✓	✓			
IFG		✓	✓	✓			
vPMC		✓	✓		✓	✓	✓
dIPMC	✓	✓	✓		✓	✓	✓
M1						✓	
IPL	✓		✓	✓	✓	✓	
SPL	✓		✓	✓	✓	✓	
STC	✓	✓	✓				✓
LOTC	✓	✓	✓	✓	✓	✓	✓
ITC	✓	✓	✓				✓
FC	✓		✓	✓	✓		✓
TP				✓			✓
LG	✓						
V1	✓						
BG						✓	
C						✓	

PFC= Prefrontal cortex; IFG=Inferior frontal gyrus; vPMC = ventral premotor cortex; dIPMC= dorsolateral premotor cortex; SPL= superior parietal lobe; IPL= inferior parietal lobe; STC= superior temporal cortex; LOTC = lateral occipitotemporal cortex; ITC= inferior temporal cortex; FC = fusiform cortex; TP= temporal pole; LG= lingual gyrus; V1= primary visual cortex; BG= basal ganglia; C=Cerebellum. ^aCulham, He, Dukelow & Verstraten, 2001, ^bGrossman, Jardin & Pyles, 2010; Saygin, 2007; Saygin, Wilson, Hagler, Bates & Sereno, 2004, ^cPeelen et al., 2013; Reynaud, Lesourd, Navarro & Osiurak, 2016; Lewis, 2006; Watson & Buxbaum, 2015, ^dRamsey, 2018; Hodzic, Kaas, Muckli, Stirn & Singer, 2009; Amoruso, Couto & Ibanez, 2011, ^eCaspers et al., 2010; Grezes & Decety, 2001; Hardwick, Caspers, Eickhoff & Swinnen, 2018, ^fHardwick et al., 2018, Gerardin et al. 2000; Orlov et al., 2010; Gallivan, McLean, Valyear, Pettypiece & Culham, 2011, ^gTurkeltaub, Eden, Jones & Zeffiro, 2002; Mechelli et al., 2005; Fedorenko & Thompson-Schill, 2014; Hagoort, 2014.

1.6 Architecture of the LOTC

To evaluate various activations for different stimuli and tasks within the LOTC, and to relate them to each other, the data for these different tasks have to be combined and compared. Combining information on different activations for different tasks examined in different people from different studies is problematic though (Brett, Johnsrude & Owen, 2002; Samartsidis, Montagna, Nichols & Johnson, 2017; Wager, Lindquist & Kaplan, 2007). Individual information from single participants is more precise in terms of the actual spatial relationship of peak responses to various tasks and their relation to individual anatomy compared to group averages. Analyses of group-based peaks might be a suitable method to detect areas of activation often related to a modular architecture of the brain where one specific brain region is active during one specific task, but peaks from different tasks are only comparable to some extent and even less so if different participants and studies have to be consulted to do so. Lingnau and Downing (2015) hypothesized that a number of selective subregions representing components of action representations like motion or bodies are part of the LOTC. Areas within the LOTC would consistently exhibit a selective response to the respective feature with possible overlaps of subareas where features are shared across domains like the movement of a hand picking up a fruit or the movement of tongs doing the same action.

Moreover, overlapping activation can contain category specific information on the representational level in terms of similarity of the underlying signal to disentangle interleaved neuronal populations (see e.g. Haxby et al., 2001; Peelen et al., 2006). Not only the strength of activity within a region but also the pattern of activation as a response to stimuli can be used to gain insights into the organization of the cortex as some representations might be more similar than others but nevertheless all can be overlapping in the same region. Information about seeing hands and moving one's own hand could be very similar in one region, while information about seeing hands and seeing moving dots might be less similar. Nevertheless, information about

seeing hands, moving the hand and seeing moving dots might be encoded in the same area by different neuronal populations, resulting in different patterns of activation.

Therefore, the aim of this project is to collect data for eight different tasks known to activate the LOTC in one study with one group of participants completing tasks and using paradigms and tasks that are as similar as possible. We use univariate analyses (individual and group GLM), multivariate analyses (RSA) as well as analyses of functional (resting state) and structural (diffusion-tensor imaging, DWI) connectivity (see e.g. Hutchison, Culham, Everling, Flanagan & Gallivan, 2014; Mars, Sallet, Schüffelgen, Jbabdi, & Rushworth, 2011) to investigate the functional and structural architecture of neuronal representations in the LOTC and further infer its function and role in the understanding of action.

The work presented here is a data-guided approach to characterize the organization and functional role of the LOTC in the context of action-related information. We aim to link findings from different domains known to be related to the processing of actions to establish a topography and examine the selectivity, sensitivity and representational similarity in LOTC and possible subregions of the area. We also validate the results of the task-based analyses with a functional connectivity-based analysis of the cortical architecture in LOTC. Only the location and overlap of single domains have been investigated before and this research program improves the direct comparison and localisation of neural representations in relation to each other by studying them in the same participants. This also enables us to evaluate the underlying organizational principle of the domains in more detail as results from functional scans of single participants, group results and resting state scans can be related to each other

The use of functional localizers as well as resting state scans gives us the opportunity to add information on the accordance of organization of LOTC derived from function and connectivity i.e. are functionally distinct regions part of distinct brain networks as well? We can further the

understanding of how different domains related to action understanding are represented in LOTC – in distinct modules or overlapping regions, represented and connected similarly or differently in the same area.

2 General methods

2.1 Participants

Twenty-one healthy adult participants (mean age: 22.95 years; age range: 19-31 years; 13 females) participated voluntarily in this study. All participants were English speaking, right-handed with normal or corrected-to-normal vision and no history of neurological or psychiatric disease. Participants were recruited using the Psychology Experiment Management System from the Department of Psychology, Royal Holloway University of London and a network of students and colleagues. Participants gave written informed consent before partaking in the study and were compensated for their time. Ethical approval was obtained from the research ethics committee of the Department of Psychology at Royal Holloway, University of London.

2.2 Experimental tasks

For this study, every participant completed eight different functional localizer tasks, a resting state and a DWI scan in two scanning sessions. In the first session four localizer tasks (high-level retinotopy, verb, basic motion and action performance), an anatomical, phase encoding scan, fieldmap and resting state scan were collected. In the second session further four localizer tasks (bodies & tools, body parts, biological motion and action observation) and the DWI scan (see table 2 for overview) were carried out. The order of the tasks was the same for all participants, shown in table 2. Exceptions from the order occurred for participant 2, where the movement performance localizer was repeated at the end of the second session due to changes in the instructions, for participant 03, where the resting state scan was repeated in session 2 because of movement in the first session, and for participant 22, who began the second session with the DWI scan due to technical problems.

Participants were familiarized with the tasks at the beginning of each session outside the scanner. They received written instructions (see Appendix A for details) for the relevant four tasks of the session and were able to ask questions about the tasks. Moreover, participants carried out a short version of each task on a PC or Laptop to ensure that they understood each task. Before each experiment, a short reminder of the task was shown inside the scanner via a Power Point presentation (one slide per task; see figures 11-18 below). After each scanning session participants filled in a short post-session questionnaire (Appendix B).

Table 2. List of Experiments, conditions, tasks, contrasts of interest for each experiment used during two sessions of MRI-scanning, the total duration of all runs together and number of runs.

Experiment	Conditions	Task	Contrast of interest	Total Duration	N runs
<i>Session 1</i>					
High Level Retinotopy	Upper Left, Lower Left, Upper Right, Lower Right	Spot change of colour of central dot	Upper Left > others Lower Left > others Upper Right > others Lower Right > others	10min 30 sec	2
Noun-Verb	Nouns, Verbs	Rate semantic similarity of two words	Verbs > Nouns	17:00 min	4
Structural Scan				6:00 min	1
Fieldmap, PE				6:00 min	1
Basic Motion	Moving Dots, Static Dots	-	Moving Dots > Static Dots	10 min30 sec	2
Action execution	Movement of Hand, Movement of Foot	Move hand or foot	Movement > Baseline	8 min30 sec	2
Resting state				10:00 min	1

Experiment	Conditions	Task	Contrast of interest	Duration	N runs
<u>Session 2</u>					
ABCT	Animals, Bodies, Chairs, Tools	One-Back Task	Bodies > Chairs Tools > Animals	21:00 min	4
Body Parts	Lower Faces, Hands, Feet, Scrambled Images	One-Back Task	All Body Parts > Scrambled Images	10 min 30 sec	2
Biological Motion	PLD of biological motion, PLD of random motion	One-Back Task	Biological Motion > Random Motion	10 min 30 sec	2
Action Observation	Action Images, Scrambled Images	One-Back Task	Action Images > Scrambled Images	8 min 30 sec	2
DWI				8 min 30 sec	1

PLD: Point-light display; DWI: Diffusion weighted imaging.

We used the SHINE toolbox in Matlab (Willenbockel et al., 2010; custom settings, matching mode= luminance; luminance option: luminance match; matching region: whole image) to control for average luminance of the images of the ABCT localizer, the body part localizer and the action observation localizer.

To create the scrambled control stimuli for the body part localizer and the action observation localizer, we used a Fourier-based decomposition technique in Matlab. In brief, each image was decomposed into a frequency and phase image using Fast Fourier Transformation (FFT). Noise was added by permuting the phase information of all pixels contained in the image. In the last step, the scrambled version of the image was reconstructed by means of inverse FFT (iFFT). In contrast to alternative strategies to add noise to an image, this fourier-based approach has the advantage that it leaves the spatial frequency content of an image intact. This is of particular

use for statistical analyses of fMRI data comparing the observation of scrambled and intact images.

2.2.1 Basic motion localizer

The localizer was adapted from Huk, Dougherty and Heeger (2002). Participants were shown alternating blocks (15s) of either white dots moving outwards on the radial axis (at a speed of 3 deg/s; number of dots: 500; limited lifetime, i.e. fraction of dots to kill on each frame: .01) or static dots (number of dots: 500; no limited lifetime, i.e. fraction of dots to kill on each frame: 0) on black background with a fixation block (15s) after each condition (example sequence in Fig. 11). Each condition block was shown five times in one run, and each run started and ended with a 15s period of fixation. Two runs of this experiment were used, each lasting 5 min 15 sec.

Task: Watch – task free

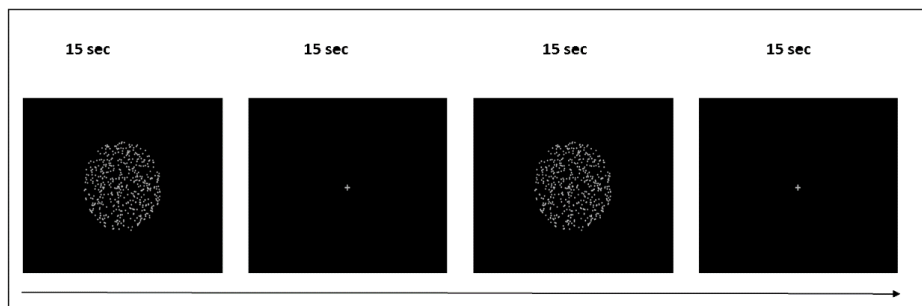


Figure 11. Example trials for basic motion localizer as shown inside the scanner right before task

2.2.2 Biological motion localizer

Blocks of animated point-light displays (PLDs) of human actions (e.g. walking, chopping or cycling or scrambled versions of the PLDs) were shown for the biological motion localizer. PLDs were used from the database available from Vanrie and Verfaillie (2004). Blocks of 5 trials per condition were used, where trials contained one animation of white dots on black background lasting 2s followed by an inter-stimulus interval of 1s. Runs consisted of five blocks per condition

interleaved with fixation blocks of 15s and additional fixation blocks at the start and end of the run, lasting 5 min 15 sec in total. Participants were instructed to press a button every time a stimulus appeared twice consecutively, which happened three times per condition in each run (example trials shown in Figure 12). The experiment consisted of two runs, each showing a total of 50 trials (25 trials per condition) using 22 different action stimuli and 22 scrambled versions. To scramble the biological motion stimuli, we applied spatial scrambling to each of the coordinates of the point light displays using the Biomotion Toolbox (van Boxtel & Lu, 2013; settings: scramble width = 50, scramble height = 100; scramble height = 50; rotation axis = x).

Task: Find same Exemplar

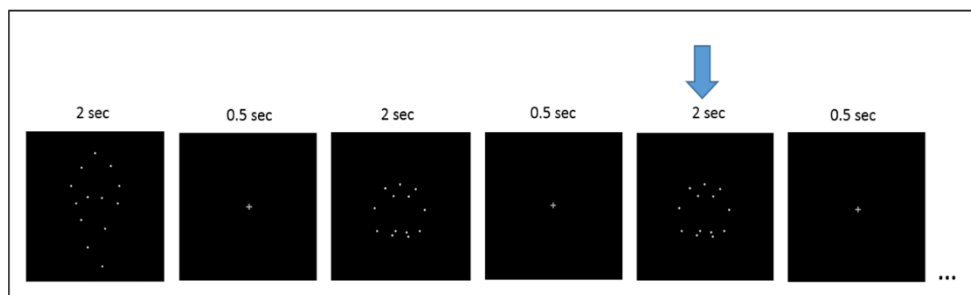


Figure 12. Example trials for biological motion localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed to detect and indicate with a button press.

2.2.3 Action execution localizer

The design of this localizer task was adopted from the motor mapping task used by Orlov, Makin and Zohary (2010). Participants were asked to move either their right hand or right foot within the scanner as if they would push a button (see Appendix A for details). The resulting movement was mostly either a tapping of the right index finger or extension of the right toe downwards. Blocks (15s) of hand and foot movement were separated by fixation blocks (15s). A movement block consisted of a colour cue (1s), indicating what body part to move next (blue circle = foot; magenta circle = hand), followed by five alternating red (1s) and green (2s) fixation crosses

prompting a break (red) or a movement of the body part (green). A total of 40 trials (20 trials per condition) divided into 8 blocks was shown in one run (example trials shown in Figure 13). Each of the two runs started and ended with 15s fixation periods and lasted 4 min 15 sec. The background colour of the screen was grey throughout the whole localizer. Blocks of motion were long enough to move the respective body part repeatedly. A small additional cushion wedge was placed under participant's right calf at the beginning of the session to allow for foot movement.

Task: Move Hand/Foot to switch on light

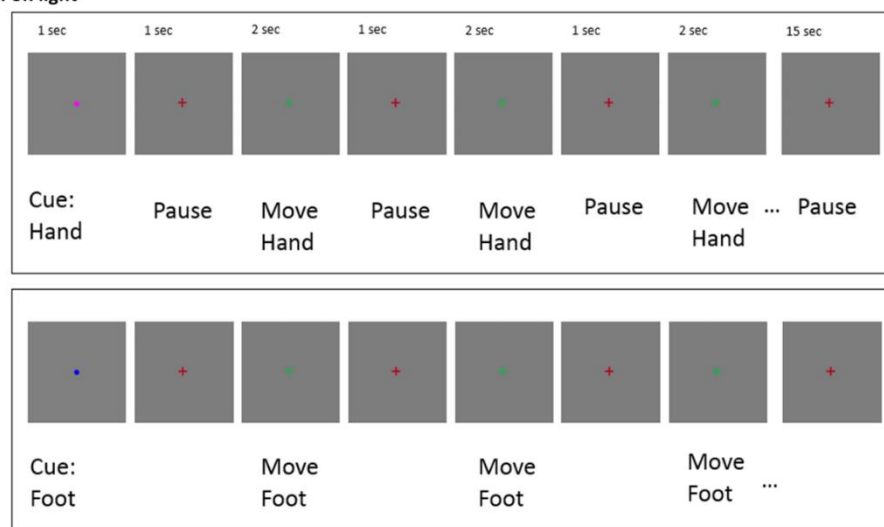


Figure 13. Example trials for movement execution localizer as shown inside the scanner right before the task.

2.2.4 Bodies & Tools localizer

The Body and Tool localizer was adapted from Downing et al. (2001). Blocks of greyscale images with white background of four conditions were used: bodies, tools, chairs and animals (Downing et al., 2001; Konkle & Caramazza, 2013). Images had a size of 300 x 300 pixels and were shown on a grey background. Blocks consisted of 15 images shown for 500 ms each with a 500 ms fixation period between images. A total of 16 blocks of visual stimulation was used per run, intermitted by a fixation block of 15s every fourth block. Moreover, each run started and ended with a fixation block of 15 s. This leads to 240 trials in total (60 per condition) for one run. Order

of blocks was counterbalanced between two versions and in each version the average position within the sequence of all condition blocks was the same. Each run lasted 5 min 15 sec and the experiment comprised four runs. As this experiment was used for two main contrasts of interest and consisted of four instead of two conditions, we used four runs. Participants had the task to indicate each repetition of the same stimulus in successive trials, occurring eight times per run (example trials shown in Figure 14).

Task: Find same Exemplar

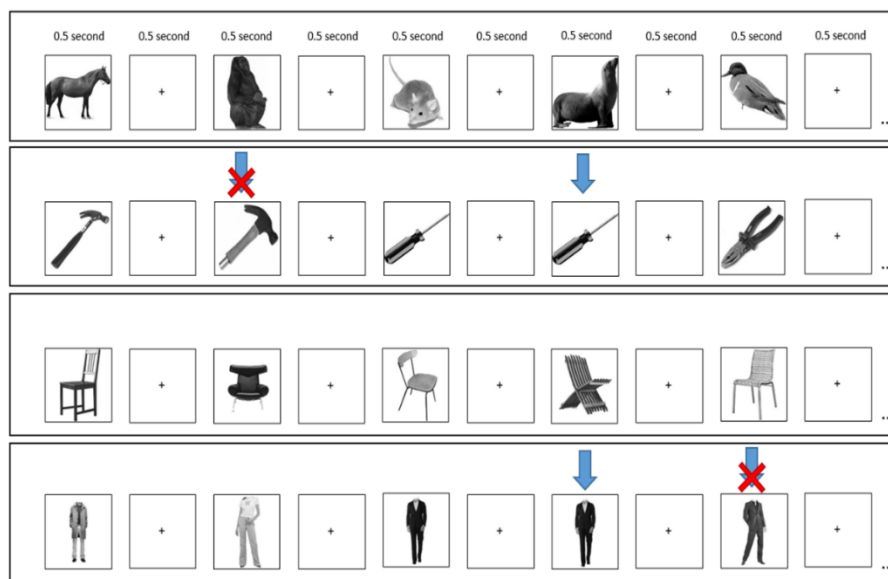


Figure 14. Example trials for Body & Tool localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed to detect and indicate with a button press while crossed out arrows show false repetitions to non-identical stimuli.

2.2.5 Body parts localizer

For the body part localizer, the paradigm was the same as the paradigm of the bodies and tools localizer, but greyscale images of hands, feet (Bracci et al., 2012), lower parts of the face (Langner et al., 2010) and scrambled images were used (60 images for each condition; example trials shown in Figure 15). Two runs of this localizer were shown with 240 trials per run (60 trials per condition).

Task: Find same Exemplar

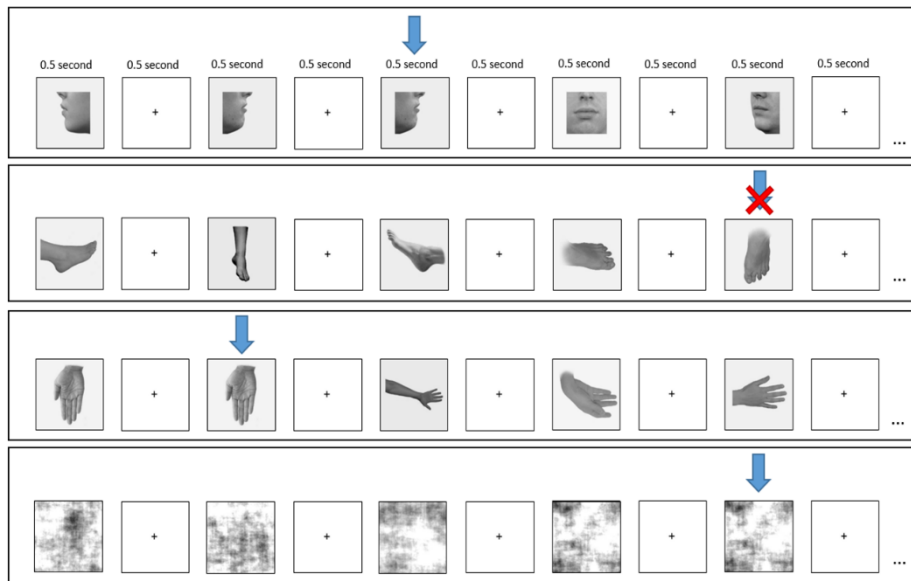


Figure 15. Example trials for body part localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed to detect and indicate with a button press while crossed out arrows show false repetitions to non-identical stimuli.

2.2.6 Action observation localizer

The action observation localizer was adapted from a task used in Tucciarelli, Wurm, Baccolo, & Lingnau (2019), showing 80 greyscale photographs of 20 different actions like riding a bike or drinking. For each action, four different images were used showing the action at different locations from a different angle with two (or four) different actors, and scrambled versions (using Fourier scrambling as described above) of these images were also created. All images were 400 x 300 pixels in size and were shown on a grey background. A block design with two conditions (action images, scrambled images) was used, alternating blocks of 10 trials consisting of an image (1s) and a fixation interval (0.5s) each with a fixation block (15s). In one run, four blocks per condition and eight fixation blocks were shown, with a total of 4 min 15 sec. Two runs were scanned for this experiment. Participants did a one-back task. For the action condition, participants were instructed to press a button every time two consecutive images showed the same action. Participants were informed that a repeated action always consisted of two

different exemplars of the same action. For the scrambled images, participants were instructed to press a button whenever the same exemplar of a scrambled action was repeated in two consecutive trials. In each run eight stimuli were repeated (example trials shown in Figure 16).

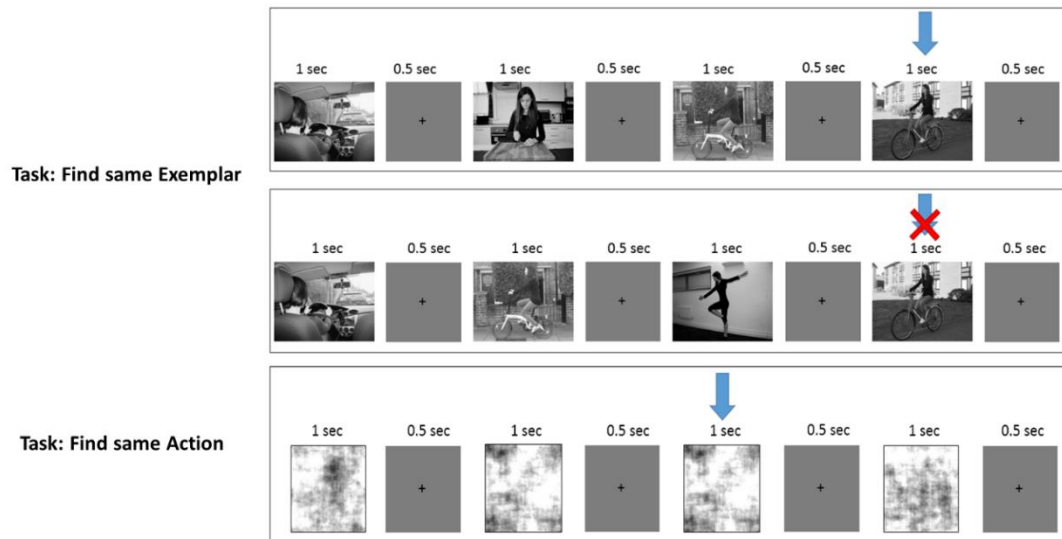


Figure 16. Example trials for action observation localizer as shown inside the scanner right before the task. Blue arrow shows a repetition of the same stimulus that participants were instructed

2.2.7 Verb localizer

The verb localizer consisted of blocks (15s) of action verbs or animal nouns (Bedny et al., 2008) alternating with fixation blocks (15s). Based on the design of Hafri, Trueswell, & Epstein (2017), within a block, for each of six trials, two words of the same category were shown on a grey screen simultaneously for 2s followed by a fixation cross for 0.5s. One word was shown above a central fixation cross, one underneath. Runs consisted of four blocks of each condition and eight fixation blocks lasting 4 min 15 sec. Each run consisted of 24 trials for each condition (48 trials overall) showing 48 words per condition. Four runs were used for this experiment. Participants were asked to rate the similarity in meaning for the two words shown together in each trial on a scale from one (not at all related) to four (highly related) by pressing one out of four buttons (example trials and button assignment shown in Figure 17).

Task: How related are the words in meaning (1-4)?

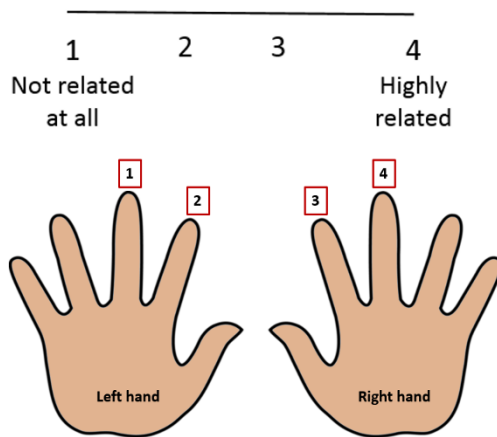
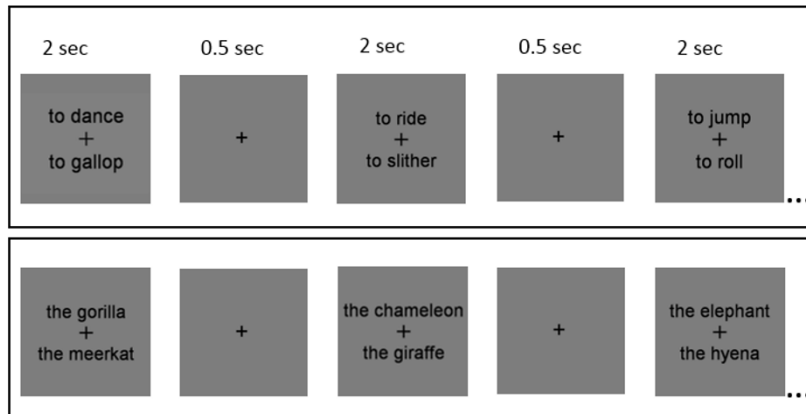


Figure 17. Example trials for Noun Verb localizer as shown inside the scanner right before the task depicting trials and assignment of responses to buttons/fingers.

2.2.8 Spatial information localizer (High Level Retinotopy)

To create a localizer for location information of natural action videos, we used a simplified experimental design adapted from Roth & Zohary (2015). For this, videos of various actions were shown in the four quadrants of the visual field. Blocks of five trials consisting of a video (2s) and a blank interval (1s) each were shown in one of the four quadrants at a time lasting 15s. Stimuli were presented in each of the four locations four times per run, while participants were asked to focus on a grey dot at the centre of the otherwise black screen. To ensure that participants kept their eyes at the centre of the screen, the central dot darkened once in every block at an

unpredictable time, and participants were instructed to press a button every time the dot got darker (example trials shown in Figure 18). Each of the 80 trials in one run showed a different video, with 20 trials in each of the four locations. The experiment consisted of two runs, lasting 5 min 15 sec each with a fixation period of 15 s at the beginning and end of the runs. This experiment allows us to investigate four visual quadrants, not more precise locations or the fovea, as this is only to get a general idea of location sensitivity, as only little detailed representation of stimulus location is expected in LOTC (Roth & Zohary, 2015).

Task: Indicate when dot briefly darkens

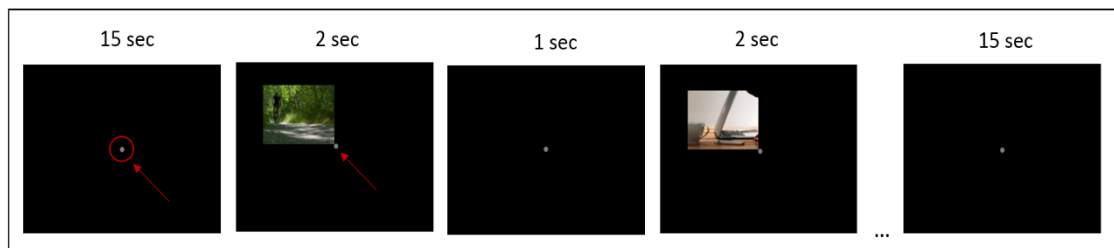


Figure 18. Example trials for High Level Retinotopy as shown inside the scanner right before task

2.3 Procedure

All stimuli were presented using Matlab R2015b (The Mathworks, Inc. Natick, USA) and ASF (Schwarzbach, 2011) utilising Psychtoolbox (Kleiner et al., 2007). Stimuli were viewed through a mirror mounted above the head coil reflecting the projection of a Sanyo, PLC-XP-100L projector (60Hz and screen resolution of 1024 x 768 pixels) at the end of the scanner bore. Responses were recorded with two button boxes in the first session and one button box in the second session. Responses were given with the right index finger, whereas for the verb localizer index and middle finger of both hands were used. Behavioural results for the different localizers are reported in Appendix C. In general, they indicate that participants properly followed the instructions.

2.4 Data acquisition

Data were acquired using a 3T Magnetom Trio (Siemens, Erlangen, Germany) scanner and a 12-channel birdcage head coil at the Combined Universities Brain Imaging Centre (CUBIC) at Royal Holloway University of London, UK. Functional images were acquired with a T_2^* -weighted gradient multi-band accelerated echo planar imaging (EPI) sequence (TR = 2 s, TE = 34 s, flip angle of 76° , a field of view of 192 mm, matrix size of 78x78 mm, voxel resolution of 2.5 x 2.5 x 2.5 mm, multiband factor = 2, 46 slices with 25% gap acquired in descending order using fat suppression). Basic Motion, Biological Motion, Body Parts, Body & Tool and High Level Retinotopy localizers consisted of 158 volumes per run and Action Observation, Action Performance and Verb localizers of 128 volumes per run. Resting-state data were collected with the same single-echo multi-band EPI sequence for 10 min (300 volumes), and participants were instructed to close their eyes.

Structural T1-weighted images were acquired with an MPRAGE sequence (TR = 1.9, TE= 3.03, flip angle of 11° , a field of view of 256 x 256, 176 slices, voxel resolution of 1 x 1 x 1 mm).

Diffusion weighted images were acquired using a pulsed-gradient-spin-echo echo-planar-imaging (PGSE-EPI) with the following acquisition parameters: TR= 8.2 s, TE = 0.09 s, 66 2 mm slices, with no gap between slices, a field of view of 192 mm, matrix size of 96 x 96 mm, voxel resolution of 2 x 2 x 2 mm.

2.5 Preprocessing for functional data

Data were preprocessed using FEAT from the FMRIB Software Library (FSL; FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl), version 5.10.0. The first 4 volumes were removed to allow stabilization of the magnetic field. Preprocessing included brain extraction using BET (Smith,

2004), coregistration and registration to MNI-standard space and the high resolution structural scan, motion correction using MCFLIRT (Jenkinson et al., 2002), high pass filtering (Gaussian-weighted least-squares straight line fitting, with $\sigma=50.0s$ /or $\text{cutoff}=100s$) and custom slice timing correction. No spatial smoothing was applied.

2.6 ROI definition

For most analyses, we focused on relevant regions of interest and used whole-brain data to obtain an overview and data from the occipito-temporal cortex or data from LOTC regions of interest defined after the multimodal cortex parcellation by Glasser et al. (2016) for more detailed analyses.

The occipito-temporal (OTC) cortex region of interest was defined using the MNI Structural Atlas. The occipital and temporal lobe regions from the atlas were combined into one OTC-ROI. Due to the lack of agreement regarding precise anatomical landmarks defining the LOTC, as it is situated at the conjunction of occipital and temporal lobe, we decided to use a large region of interest first for the localisation of peak responses.

To avoid the definition of ROIs based on our functional data or not agreed on anatomical markers, more specific and smaller regions of interest (ROIs) were based on an independent brain map - the multimodal cortex parcellation from Glasser et al. (2016) (Human Connectome Project's multi-modal parcellation, version 1.0 (HCP_MMP1.0)). This parcellation is based on independent data including structural, functional and connectivity measurements from the human connectome project and can be converted to individual subject space. The multi-modal parcellation HCP_MMP1.0 is based on measurements of relative cortical myelin content and cortical thickness to evaluate structural measures, seven different functional localizers (working

memory, gambling, motor, language, social cognition, relational processing and emotional processing task) for cortical function and resting-state functional MRI revealing connectivity and topography in the cortex.

For each participant, the segments of the LOTC-ROI are defined by the following steps. First, the anatomy of each participant is segmented using FreeSurfer. Next, the Glasser parcellation is applied to these individual anatomies, generating FreeSurfer annotation files projected on a template brain (fsaverage). Next, these annotation files are transformed to each individual subject's space and converted to volume masks. The ROI created by this pipeline consists of the cortical areas Area Lateral Occipital 1 (LO1), Area Lateral Occipital 2 (LO2), Area Lateral Occipital 3 (LO3), Visual area 4 temporal (V4t), Middle Temporal Area (MT), Medial Superior Temporal Cortex (MST), fundus of the STS area (FST), Basal parietal area (PH), Basal (temporooccipital) parietal area at temporal entrance (PHT), Middle temporal area proper posterior (TE1p), Superior Temporal Sulcus ventral posterior (STSvp) as defined in Glasser et al. (2016) for both hemispheres separately (see Figure 19). Regions were defined for each individual participant and for the MNI standard brain to use for further analyses.

3 Spatial organization in LOTC

3.1 Introduction

The first questions to examine are where in the brain regions responding to different localizer tasks are located, and how the locations of these regions relate to each other spatially. This enables us to replicate the findings summarized in Chapter 1, and to verify that the selected localizers indeed engage regions in LOTC when focusing on the response magnitude. To identify brain regions recruited by the different functional localizers, a univariate analysis with a general linear model is applicable (Smith, 2004; Kriegeskorte & Bandettini, 2007). A further question is concerned with how strongly a given region of interest responds to the various functional localizers, and the degree to which a given region of interest shows a preference for one of the functional localizers. Hence, in this chapter we use univariate analyses to investigate the activation and organization of the LOTC at the level of the OTC as a whole, and at the level of regions of interest identified based on the Glasser parcellation described in the previous chapter.

3.2 Analyses

3.2.1 Neural responses for different domains in the brain

The replication of previous findings of activation within the LOTC region for the localizer tasks basic motion, biological motion, tools, bodies, body parts, action observation, action performance and verbs was the first objective. To this aim, per participant, session, run and localizer experiment, a GLM-based analysis with a design matrix containing condition block predictors convolved with a gamma function and 6 motion parameters from motion correction was performed for the whole brain. Data from separate runs were combined using a second

level fixed effects analysis and a random-effects (RFX) group level GLM analysis was conducted for each experiment. To correct for multiple comparisons, nonparametric permutation testing (randomise with 10000 permutations) was used applying threshold free cluster enhancement with $p < 0.05$ using FSL's randomise (Winkler, Ridgway, Webster, Smith & Nichols, 2014). Threshold free cluster enhancement was chosen to correct for multiple comparisons as it was shown to increase sensitivity for the signal and gives more interpretable results while using less arbitrary settings like the threshold of clusters (Smith & Nichols, 2009).

For participant 1 the action observation, body part and biological motion localizer and for participant 17 the second run of the biological motion localizer were not included in the group analysis due to excessive head motion. Runs showing repeated displacement of the head combined with displacement $> 2\text{mm}$ were excluded after inspection. The main RFX GLM-contrasts of interest for this work are basic visual motion of dots versus static dots (basic motion), biologically moving point light displays versus randomly moving PLDs (biological motion), pictures of tools versus pictures of animals (tools), pictures of body parts (faces, hands and feet) versus scrambled pictures (body parts), pictures of bodies versus pictures of chairs (bodies), pictures of actions versus scrambled pictures (action observation), hand and foot action performance versus baseline (action performance) and verbs versus nouns (verb) (see also Table 2).

To augment the meta-analytic findings of the different domains activating the same area as shown in the introduction (Figure 9), the single results from the contrasts for basic visual motion, biological motion, tools, bodies, body parts, action observation, action performance and verbs were combined to display overlapping activated regions. Moreover, from this map we calculated the percentage of overlap of activated (=nonzero) voxels.

Results were mapped to a freesurfer surface using MNI2fsaverage (Wu et al., 2018), and the annotation files of the multimodal parcellation of the cortex HCP-MMP1.0 by Glasser et al. (2016) was used to show boundaries of different brain areas.

To evaluate the spatial overlap of different functional localizers, the resulting p-maps were used to threshold results at $p < 0.05$. Next, maps were binarized, and a unique value was assigned to each contrast of interest. The maps were then combined to calculate the percentage of overlapping areas and discern combinations of contrasts. For visualisation, the binarized maps were combined to show the absolute number of overlapping contrasts. To give an overview of combinations of contrasts in the same voxel, a summary table was created for the voxels part of five, six or seven activation maps from localizers.

3.2.2 Topography of peak activation

As we are interested not only in the average location but also the topographic location of different domain-activations for all participants in relation to each other, we used peak-t values for each contrast of interest, separately for each participant, for this analysis. To include only voxels in and around the LOTC, results from the whole brain GLM analysis were masked with an occipital and temporal lobe (OTC) mask for the left and right hemisphere created using the MNI Structural Atlas (see section 2.6 for details).

To investigate whether the location of peaks resulting from the contrast of interest is consistent across participants, we evaluated the peaks of each contrast using the highest t-value. The location of the maximum t-value of each contrast of interest from each participant within occipital-temporal cortex was extracted using `fsl featquery`. A 95%-confidence interval for the mean location across participants for each contrast of interest was calculated for each dimension (x, y, z) and informs the x, y and z dimension of an ellipsoid depiction of the interval.

3.2.3 *Selectivity for domains in LOTC subparcels*

To depict how selective LOTC-subregions defined after HCP_MMP1.0 by Glasser et al. (2016) in standard MNI-space are for single contrasts, we extracted the average t-value for each area from the individual GLM results for all contrasts of interest and calculated the average across all participants. The resulting average t-values from each domain in each region were then ordered in a descending manner and depicted in bar-charts. Highly selective regions would be expected to exhibit strong responses (high t-values) to one domain but not to other domains (low t-values) while a region with low selectivity would show strong responses to several domains or none. This would result either in an abrupt drop of the t-score for domains the region is not selective for, or stable or a gradually decreasing t-value if the region is not selective. To test for individual effects of the eight different domains in each region, t-tests were calculated with a threshold of $p < 0.01$ (two-tailed).

3.3 Results

3.3.1 *Neural responses for different domains in the brain*

First, we wanted to identify the regions revealed by the different RFX GLM contrasts of interest. To replicate previous studies, whole brain analyses on the group level were computed for each contrast and mapped to the surface. The resulting maps are shown in Figures 20, 21 and 22.

For all eight contrasts of interest, responses in the left LOTC were found with contrasts resulting in responses covering extensive parts of LOTC, in particular for action observation, basic motion and body parts, while tools and bodies show smaller clusters. The responses for biological motion, action performance and verbs were relatively focal in LOTC (Figure 20 and 21, A-H). For

right LOTC, comparable responses were found except for tools, action performance and verbs, which did not activate significant clusters in right LOTC (Figure 22, A-H).

The RFX GLM contrast basic visual motion > static revealed clusters in occipital cortex, LOTC and also small clusters in fusiform cortex (Figures 20-22, panel A). This was found similarly in the left and right hemisphere.

Responses to the biological motion contrast were restricted to LOTC and fusiform cortex and clusters were quite localized (Figures 20-22, panel B).

The RFX GLM contrast for tools showed revealed clusters in the left hemisphere in lateral and superior regions of occipital cortex, extending to LOTC and superior and inferior parietal lobe and a small region in dlPMC (Figures 20-21, C). In the right hemisphere, only superior and inferior parietal lobe showed significant responses (Figure 22, panel C).

Responses for the contrast bodies > chairs were found in LOTC, superior parietal lobe (SPL), inferior frontal gyrus (IFG) and fusiform cortex (FC) for both hemispheres but with smaller clusters in the left hemisphere and extensive clusters in the right hemisphere (Figures 20-22, panel D). Additionally, IPL and vPMC were activated in the right hemisphere.

Areas responding more during the observation of body parts compared to scrambled images encompass LOTC, fusiform cortex and inferior and superior parts of occipital cortex in both hemispheres (Figures 20-22, panel E).

As expected for movements with the right hand and foot, the activation obtained during action performance compared to baseline was localized more strongly in the left hemisphere including LOTC, dorsolateral and ventral premotor cortex, inferior frontal gyrus, inferior and superior parietal cortex (Figures 20-21, panel F). Responses were weaker in the right hemisphere and

were located in the occipital pole, angular gyrus, ventral premotor cortex and inferior parietal lobe (Figure 22, F).

Responses to the action observation contrast covered extensive areas around LOTC from occipital pole, superior occipital areas, posterior regions of superior temporal gyrus and central regions of inferior temporal cortex. Additionally, regions in fusiform cortex, inferior frontal gyrus, prefrontal cortex and the temporal pole were found (Figures 20-22, panel G). Similar areas were recruited in the left and right hemisphere, but right frontal areas were considerably less pronounced.

For verbs > nouns responses were observed in anterior and superior regions of LOTC and in inferior frontal cortex in the left hemisphere (Figures 20-21, panel H). No cluster in the right hemisphere survived correction for multiple comparisons (Figure 22, H).

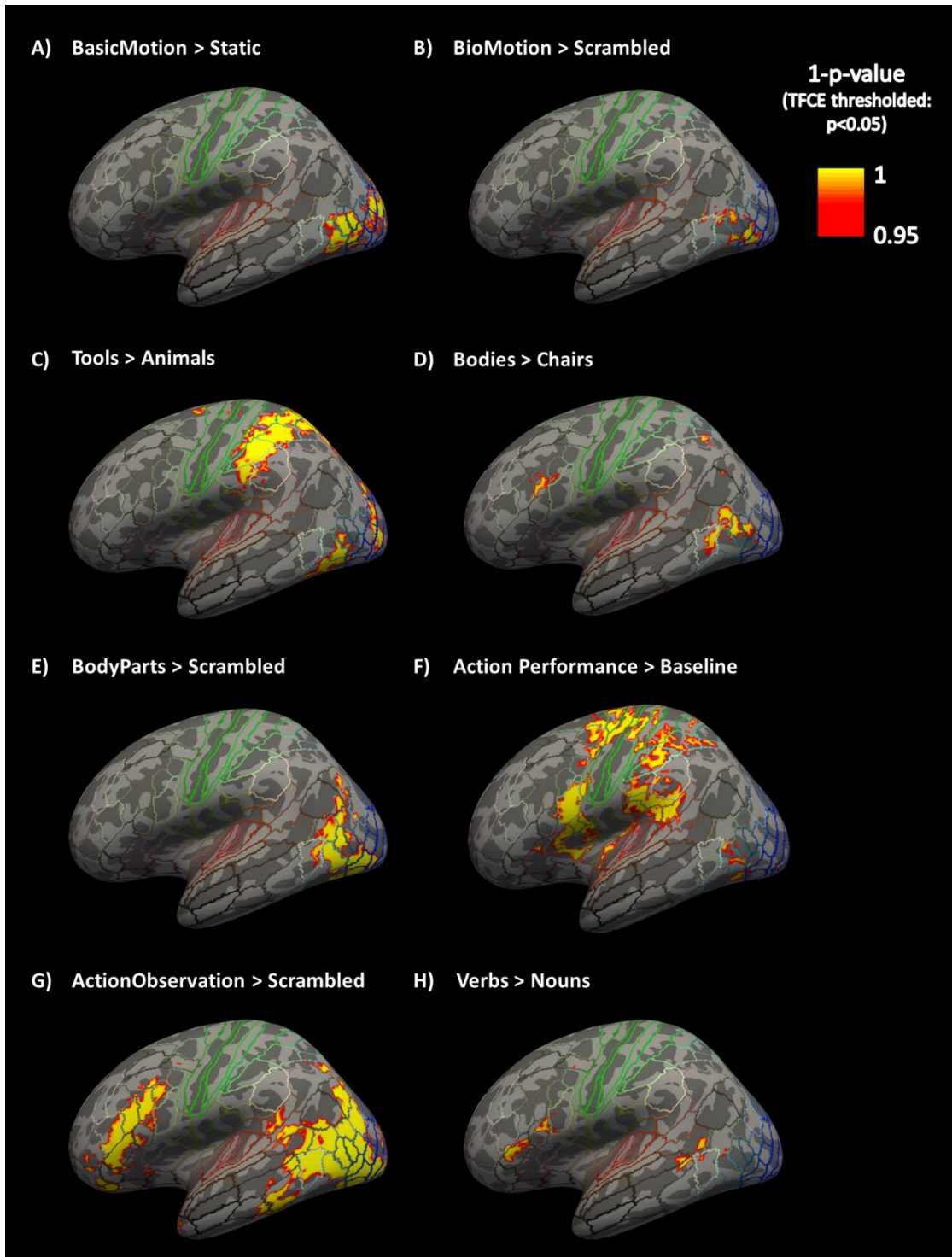


Figure 20. Left lateral view of results of the univariate group level GLMX-analysis (with $p < 0.05$, Threshold free cluster enhancement (TFCE) corrected) for eight main contrasts of interest with overlaid cortex parcellation from the human connectome project (HCP-MMP1.0).

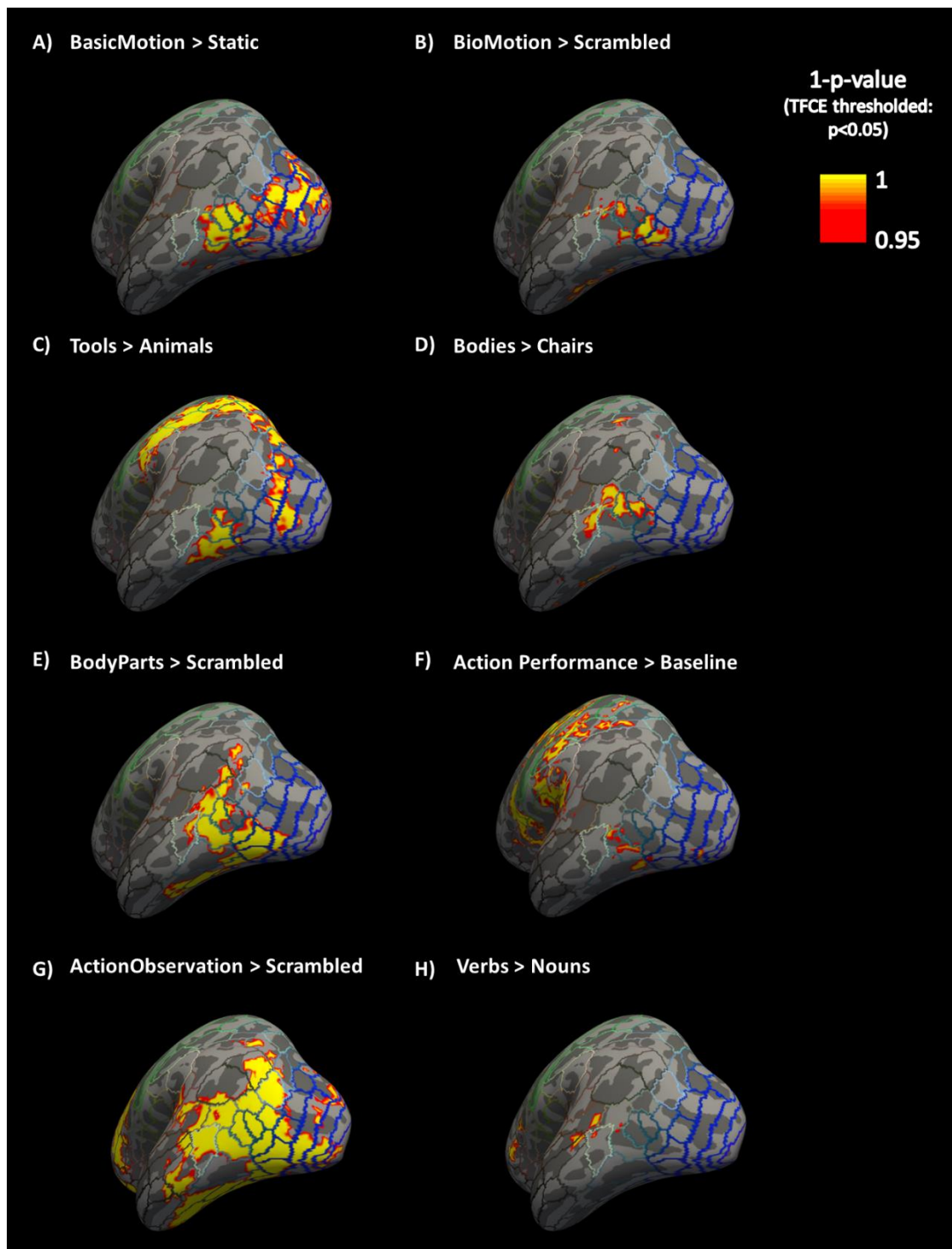


Figure 21. Left posterior-lateral view of results of the univariate group level GLMX-analysis (with $p < 0.05$, Threshold free cluster enhancement (TFCE) corrected) for eight main contrasts of interest with overlaid cortex parcellation from the human connectome project (HCP-MMP1.0).

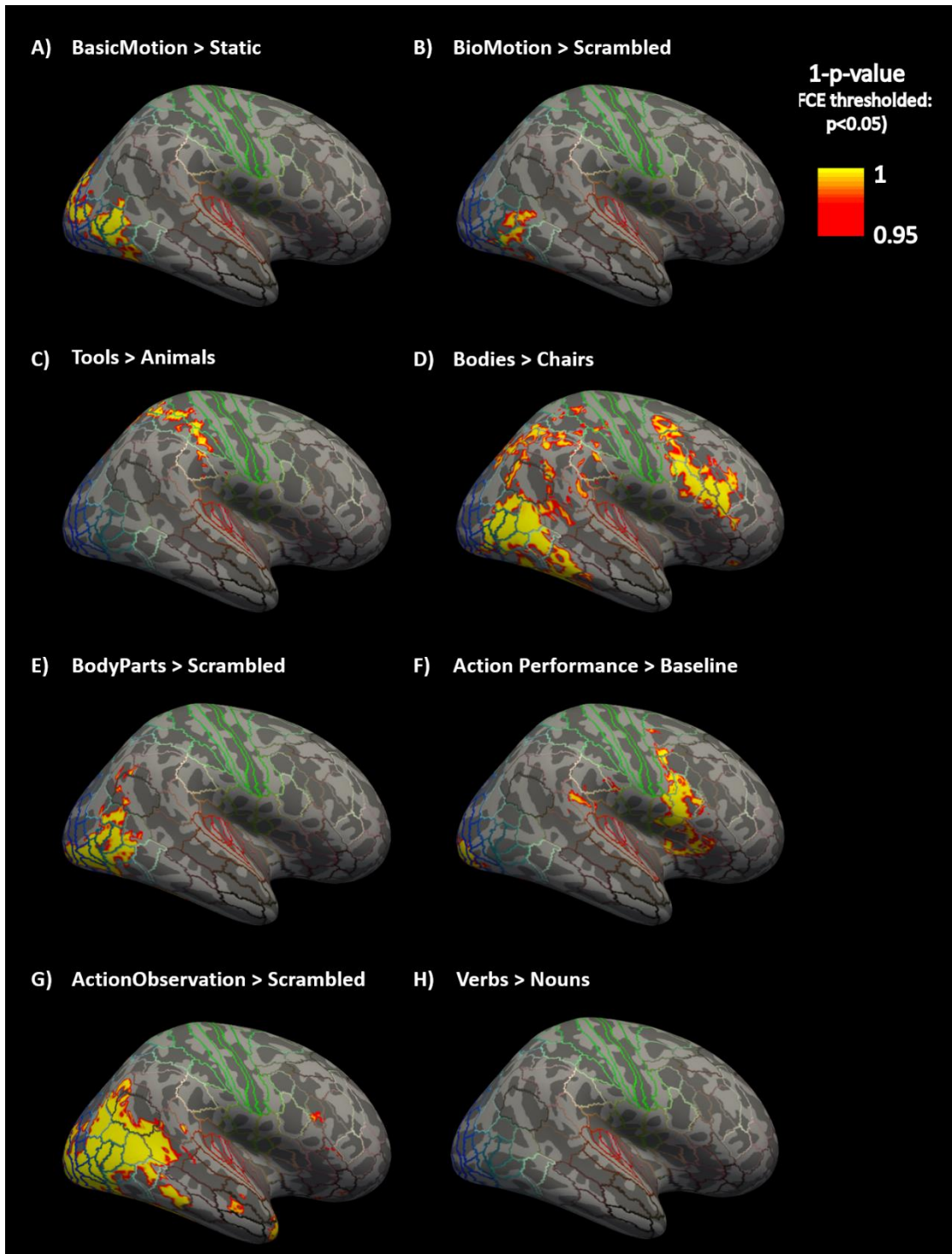


Figure 22. Right lateral view of results of the univariate group level GLMX-analysis (with $p < 0.05$, Threshold free cluster enhancement (TFCE) corrected) for eight main contrasts of interest with overlaid cortex parcellation from the human connectome project (HCP-MMP1.0).

Next, we aimed to determine overlapping responses, i.e. for voxels that show significant activation for more than one contrast, at the group level (see section 3.2.1 for details). Overlapping voxels were primarily located in the LOTC region and in the fusiform cortex and out of the eight contrasts up to seven showed responses in the same voxels (Figure 22).

All RFX GLM contrasts revealed overlapping voxels (see Table 3 for overview for highest numbers of overlapping contrast) and out of a total of 69981 activated voxels overall, 13.52% were activated by two different localizers, 3.47% by three, 1.51% by four, 0.49% by five, 0.09% by six and 0.009% by seven contrasts of interest from the eight localizers. No clear pattern of coactivation was found, but the most frequently overlapping contrasts are basic motion, biological motion, body parts, bodies and action observation occurring together in 234 voxels.

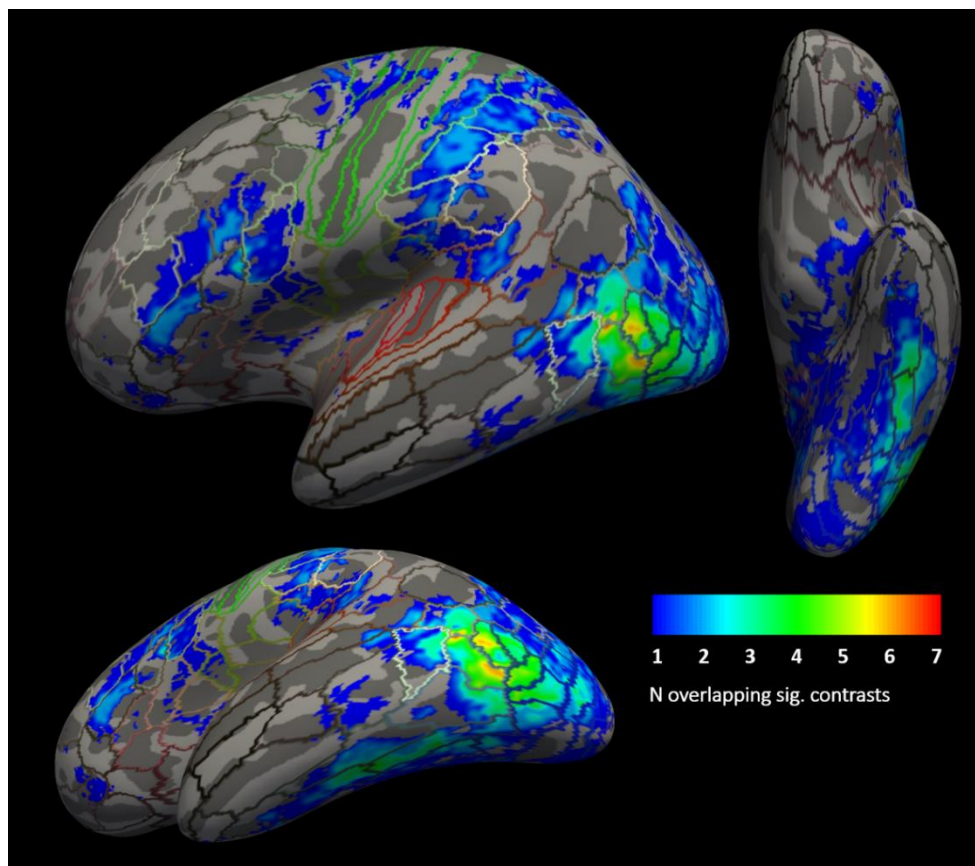


Figure 22. Number of contrasts showing significantly activated regions from univariate group level GLMX-analysis (with $p < 0.05$, Threshold free cluster enhancement (TFCE) corrected) from eight main contrasts of interest with overlaid cortex parcellation from the human connectome project (HCP-MMP1.0).

Table 3. Contrasts overlapping in voxels with 5, 6 or 7 significantly active contrasts.

N contrasts overlap	N of Voxels	Basic Motion	Biological Motion	Tools	BP	Bodies	Action Observation	Action Performance	Verbs	
<u>5</u>	234	✓	✓		✓	✓	✓			
	18	✓			✓	✓	✓	✓		
	17	✓		✓	✓		✓	✓		
	16	✓	✓		✓		✓	✓		
	15	✓	✓	✓	✓		✓			
	14		✓		✓	✓	✓	✓		
	10	✓		✓	✓	✓	✓			
	6	✓	✓				✓	✓		
	6		✓			✓	✓	✓		
	5	✓	✓	✓	✓				✓	
	2					✓	✓	✓	✓	✓
	<u>6</u>	27	✓	✓		✓	✓	✓	✓	
14		✓	✓	✓	✓		✓	✓		
6		✓		✓	✓	✓	✓	✓		
5		✓	✓	✓	✓	✓	✓			
5			✓		✓	✓	✓	✓	✓	
4		✓	✓		✓		✓	✓	✓	
2		✓	✓			✓	✓	✓	✓	
1		✓	✓		✓	✓	✓	✓		
<u>7</u>	5	✓	✓	✓	✓	✓	✓	✓		
	1	✓	✓		✓	✓	✓	✓	✓	

3.3.2 *Topography of peak activation*

To evaluate the topographic array of the contrasts of interest, we extracted the location of individual's peak t-values of all RFX GLM contrasts of interest. These are shown in MNI space in Figure 23. The peak activation scatter plots show that most tasks have a similar location across participants, but some are more consistent than others. The contrast for verbs (red) is located most anterior in comparison to the other contrasts for most participants and the confidence interval is small in comparison to, for example, the contrast of biological motion (blue). Going from anterior to posterior, the next cluster of peaks is from the action performance (orange) experiment with a wider spread of peaks. The contrast for bodies showed peak activations posterior to the peaks revealed for Action Performance, but the smaller confidence interval reflects the limitation of individual peak locations to the occipital area. In most posterior areas, activation by bodies (light green), motion (purple), tools (light blue), biological motion and body parts (dark green) clustered together, listed from the most superior cluster to the more inferior clusters. Located more anterior and more inferior, peak t-values from the Action Observation (yellow) experiment were clustered. Biological motion and basic motion peaks show less consistency regarding the location across participants, resulting in bigger confidence intervals compared to the more closely spaced peaks of the functional localizers for bodies, tools and body parts. The cluster for action observation is slightly offset with peaks located more anterior and the location is very similar across participants.

The overall topography across both hemispheres is about the same, but although the mean location of the clusters in relation to each other are very similar, the spread of individual peaks is higher in the right hemisphere compared to the left.

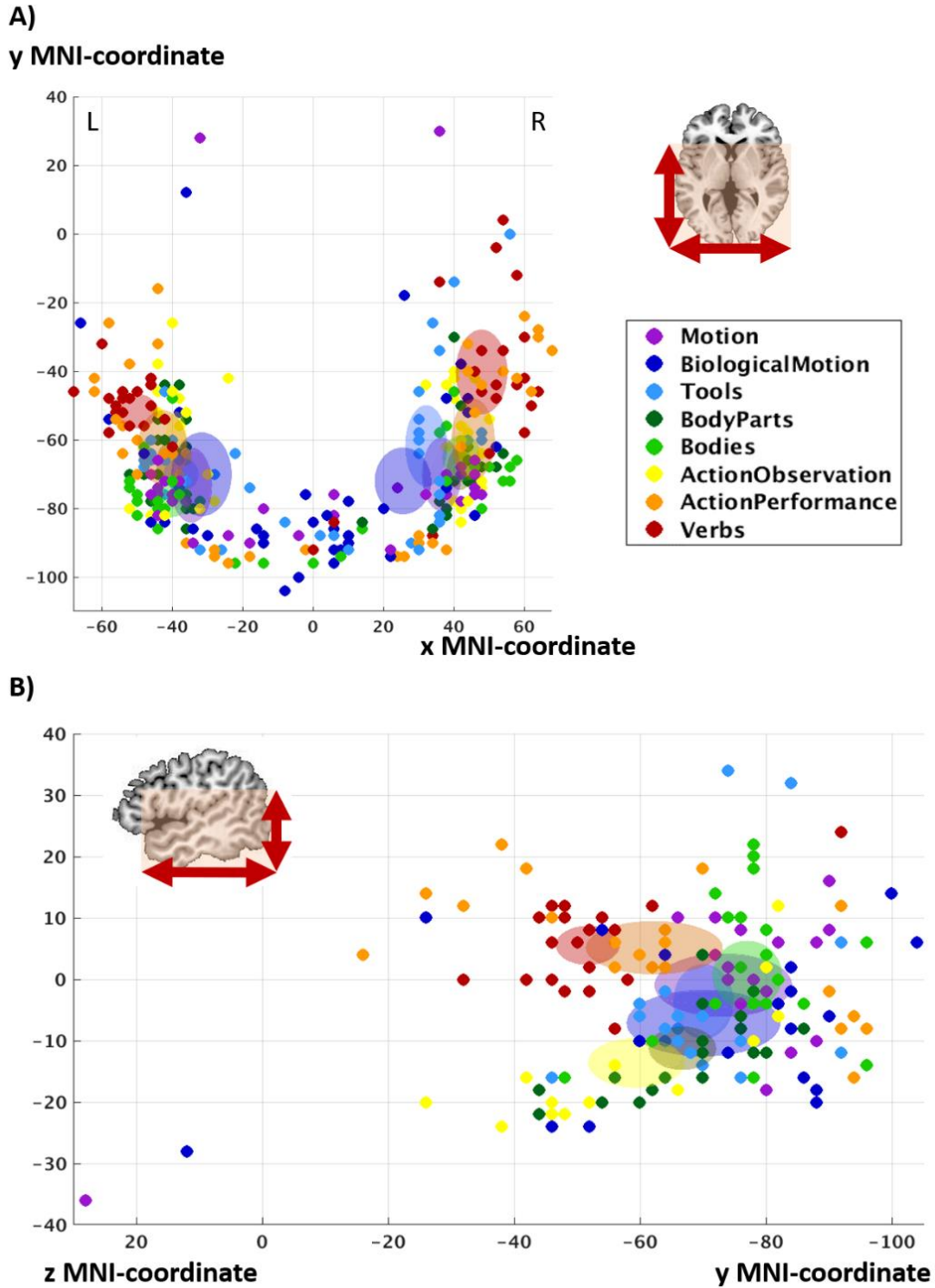


Figure 23. MNI-space location of single subject peak-t-scores for each contrast of interest in occipito-temporal cortex. Transparent ellipses depict 95%-confidence interval around the average location for the respective contrast A) 3D coordinates seen from horizontal plane, B) 3D coordinates seen from sagittal plane.

3.3.3 *Selectivity of subareas*

To investigate modularity of subparcels of the LOTC, we used the multimodal cortex parcellation HCP_MMP1.0 and chose 11 regions of interest in both hemispheres (Figures 24 and 25) in which we determined how selectively active they were for the contrasts of interest. The mean t-value from all contrasts in each of the regions of interest were calculated and ordered in a descending manner (Figure 25, panel A). This indicated that values in some regions show a subtle decrease across the different contrasts in LO1, MT, MST, TE1p or STSvp or responses were of similar strength across several contrasts in LO2, FST, PH and PHT and in some regions single contrasts stand out like LO2, LO3, PH in the left hemisphere. In the right hemisphere (see Figure 25, panel B) a decrease across contrasts can be seen in most of the regions with some plateaus for example in LO1, PH and only very little differences for several contrasts, likely due to small overall t-values, in PHT, TE1p and STSvp.

To check for effects of different contrasts in the subregions of LOTC, we used one-sample t-tests and the detailed results can be found in table 4 (left hemisphere) and table 5 (right hemisphere). In the left hemisphere, action observation showed a selective effect in LO1, and in the more anterior LO2 effects for action observation, body parts, basic motion and biological motion were found. Region LO3 responded most strongly to action observation but also to bodies, body parts and the contrasts for action performance and tools yielded negative effects in LO3. More anterior V4t and MT revealed strong effects for action observation, body parts, basic motion, biological motion and bodies and additionally in MT also an effect for action performance. MST showed the strongest effect for basic motion followed by action observation, tools, body parts, biological motion and bodies while the more anterior FST responded strongly to tools, action observation, body parts and action performance, basic motion and biological motion also activated FST while the subregion PH showed a selective response to the contrast of tools. PHT at the same time showed significant effects for action observation, verbs and bodies and the

most anterior regions TE1p and STSvp responded to action observation and negatively to action performance, and an effect was found in TE1p for bodies and in STSvp the strongest effect was for verbs.

Overall, for most regions (7 out of 11) in the left LOTC areas, responses to the contrast of action observation were the strongest. Only in LO1 was action observation the sole significant contrast, while in the other regions at least two other contrasts activated the area significantly. For MST, basic motion was the strongest contrast and in FST tools, but both regions responded to five other contrasts also significantly and only in PH was tools the strongest and only significant contrast.

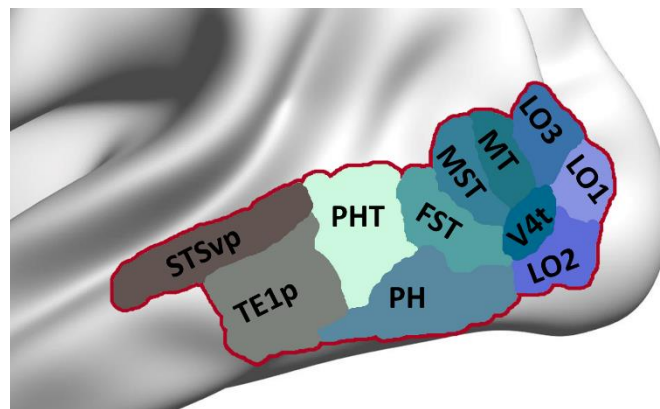


Figure 24. Left hemisphere regions of interest in LOTC area defined after the HCP-MMP1.0 cortex parcellation by Glasser et al. (2016)

Contrast

ROI	Biological Motion															
	Motion		Tools		Body Parts		Bodies		Action Observation		Action Performance		Verbs			
	t(20)	p	t(19)	p	t(20)	p	t(19)	p	t(20)	p	t(19)	p	t(20)	p		
L_LO1	2.465	0.023	-1.653	0.114	1.289	0.212	-2.726	0.013	-2.780	0.012	3.462	<0.01	-2.660	0.015	-1.411	0.174
L_LO2	3.608	<0.01	3.856	<10 ⁻⁴	-1.222	0.236	7.698	<10 ⁻⁷	1.952	0.065	6.049	<10 ⁻⁶	2.759	0.012	-0.336	0.740
L_LO3	2.765	0.012	1.587	0.128	-3.027	<0.01	3.430	<0.01	3.620	<0.01	6.158	<10 ⁻⁶	-3.557	<0.01	-2.551	0.019
L_V4t	6.080	<10 ⁻⁶	5.283	<10 ⁻⁵	1.998	0.059	4.695	<10 ⁻⁴	3.183	<0.01	7.667	<10 ⁻⁷	1.730	0.099	0.531	0.601
L_MT	5.678	<10 ⁻⁵	5.150	<10 ⁻⁵	0.972	0.343	5.022	<10 ⁻⁵	5.809	<10 ⁻⁵	5.906	<10 ⁻⁶	2.118	0.047	1.312	0.204
L_MST	6.029	<10 ⁻⁶	4.227	<10 ⁻⁴	4.095	<10 ⁻⁴	4.389	<10 ⁻⁴	3.292	<0.01	5.322	<10 ⁻⁵	2.691	0.014	-0.348	0.731
L_FST	4.047	<10 ⁻⁴	3.760	<0.01	4.745	<10 ⁻⁴	4.274	<10 ⁻⁴	0.663	0.515	4.318	<10 ⁻⁴	3.446	<0.01	1.804	0.086
L_PH	2.116	0.047	1.715	0.102	5.677	<10 ⁻⁵	2.052	0.054	-1.734	0.098	1.796	0.088	1.667	0.111	-1.718	0.101
L_PHT	-1.149	0.264	2.464	0.023	0.937	0.360	2.524	0.020	3.895	<10 ⁻⁴	6.945	<10 ⁻⁷	1.469	0.157	4.376	<10 ⁻⁴
L_TE1p	-1.588	0.128	0.631	0.535	-0.939	0.359	0.518	0.610	2.984	<0.01	4.508	<10 ⁻⁴	-3.131	<0.01	-0.210	0.836
L_STSvp	-2.140	0.045	0.349	0.731	0.026	0.980	0.846	0.408	1.356	0.190	2.860	<0.01	-3.592	<0.01	3.285	<0.01

For the right hemisphere, similar effects were found with strongest responses to action observation in eight out of eleven regions (Figures 26 and 25, panel A). Additionally, the order of the other contrasts showed some consistency with body parts, bodies and basic motion in second, third and fourth place in different configurations followed by biological motion with the fifth strongest effect and action performance, verbs and tools with the remaining positions in seven subregions. LO1 reveals significant responses only to action observation and basic motion, while LO2 responds to action observation, body parts, bodies, biological motion and negatively to tools. In LO3 the strongest effect was found for action observation and a smaller effect also for basic motion. For the central regions V4t, MT, MST, FST and PH significant responses were found for the contrasts action observation, body parts, bodies and motion and individually V4t also responded to biological motion and tools, MT to tools, FST to biological motion and action observation and PH to biological motion. The more anterior regions PHT, TE1p and STSvp effect were smaller and the strongest contrast was bodies, but only significantly activated in PHT and TE1p. PHT also showed an effect for action observation, STSvp responded selectively to action performance.

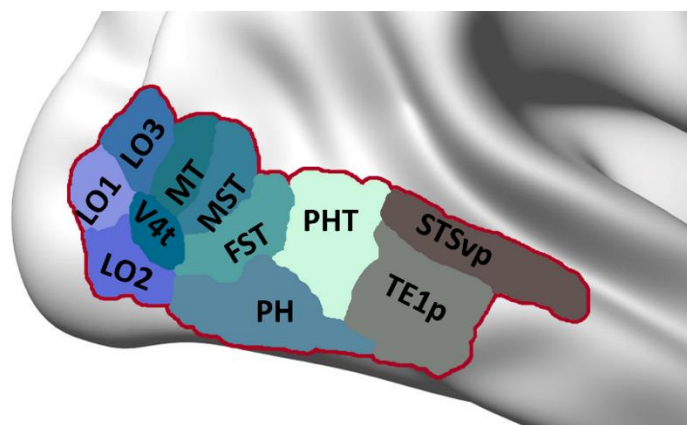


Figure 26. Right hemisphere regions of interest in LOTC area defined after HCP-MMP1.0 cortex parcellation by Glasser et al. 2016.

Table 5. Results for one-sample t-test for effects (uncorrected for multiple comparisons) of all contrasts in the regions of interest in the right hemisphere. T-values were extracted for each individual participant and region of interest (ROI) and averaged across participants for a mean t-value per contrast and region.

ROI	Contrast															
	Motion		Biological Motion		Tools		BodyParts		Bodies		Action Observation		Action Performance		Verbs	
	t(20)	p	t(19)	p	t(20)	p	t(19)	p	t(20)	p	t(19)	p	t(20)	p	t(20)	p
R_LO1	4.766	<10 ⁻⁴	1.561	0.134	1.897	0.072	-0.121	0.905	-2.672	0.015	5.145	<10 ⁻⁵	-0.263	0.795	-0.376	0.711
R_LO2	2.231	0.037	4.809	<10 ⁻⁴	-5.160	<10 ⁻⁵	6.742	<10 ⁻⁶	4.454	<10 ⁻⁴	8.279	<10 ⁻⁸	1.751	0.095	1.035	0.313
R_LO3	3.180	<0.01	1.730	0.099	-1.325	0.200	1.812	0.085	1.203	0.243	6.102	<10 ⁻⁶	-2.633	0.016	-0.840	0.411
R_V4t	2.955	<0.01	5.302	<10 ⁻⁵	-5.175	<10 ⁻⁵	6.440	<10 ⁻⁶	4.875	<10 ⁻⁵	10.269	<10 ⁻⁹	1.701	0.104	0.917	0.370
R_MT	4.862	<10 ⁻⁵	2.269	0.035	-4.256	<10 ⁻⁴	5.499	<10 ⁻⁵	8.283	<10 ⁻⁸	8.072	<10 ⁻⁷	-0.297	0.770	0.462	0.649
R_MST	5.341	<10 ⁻⁵	2.450	0.024	-0.085	0.933	3.502	<0.01	5.322	<10 ⁻⁵	3.643	<0.01	0.672	0.510	-1.110	0.280
R_FST	5.361	<10 ⁻⁵	4.398	<10 ⁻⁴	1.458	0.160	5.590	<10 ⁻⁵	6.093	<10 ⁻⁶	4.997	<10 ⁻⁵	3.250	<0.01	-0.951	0.353
R_PH	3.416	<0.01	2.900	<0.01	1.371	0.186	4.086	<10 ⁻⁴	5.249	<10 ⁻⁵	3.164	<0.01	1.529	0.142	-1.272	0.218
R_PHT	-0.022	0.983	1.333	0.198	0.278	0.784	0.856	0.402	8.029	<10 ⁻⁷	2.851	<0.01	1.226	0.234	0.342	0.736
R_TE1p	-0.621	0.542	-0.207	0.838	-1.419	0.171	-1.738	0.098	5.248	<10 ⁻⁵	2.554	0.019	-2.244	0.036	-2.392	0.027
R_STSvp	-2.267	0.035	-2.540	0.019	0.117	0.908	-1.375	0.184	1.670	0.111	0.243	0.810	-3.632	<0.01	-1.870	0.076

3.4 Discussion

The results revealed by the analysis of the spatial arrangement and the selectivity of the different functional localizers show the engagement of LOTC in different action related domains with a topography of the domains and overlapping responses in central LOTC. The investigation of possible subregions showed that focal regions exhibit no selectivity for the tested contrasts while some subregions along the edge of our LOTC responded selectively to one or only few contrasts.

First, we examined responses of the LOTC area and subregions to action related contrasts and found significant responses for basic motion, biological motion, tools, body parts, bodies, action observation, action performance and verbs in the left hemisphere LOTC. The overlap of up to seven contrasts in the same voxels in LOTC indicate it is a hub region for several domains. The underlying function and organization might not be a small-scale modular arrangement (Fodor, 1983; Fodor, 2001; Caramazza & Shelton, 1998; Kanwisher, 2010) of single domains as we found little indication of functional selectivity and isolation of regions but rather extensive overlap and distribution of responses across the LOTC that could indicate a representational space with overlapping map representing different features of the used stimuli (Op de Beeck, Haushofer & Kanwisher, 2008, Aflalo & Graziano, 2006) .

We replicated the previously reported activation of LOTC areas for basic motion (e.g. Papeo & Lingnau, 2015), biological motion (e.g. Grossman et al., 2000), tools (e.g. Bracci et al., 2011), body parts (e.g. Orlov et al., 2010), bodies (e.g. Downing et al, 2001), action observation (e.g. Caspers et al., 2010), action performance (e.g. Orlov et al., 2010) and verbs (e.g. Bedny et al., 2008) and could show that not different, clearly separable clusters responded to the different domains but at the group level clusters in the centre of LOTC showed overlapping responses for all contrasts like partially overlapping modules suggested by Lingnau and Downing (2015). The

left lateralisation is in line with the well documented left hemispheric language responses (Knecht et al., 2000) and the left lateralized motor response for the action performance, as all participants were using their right hand and foot and therefore recruited left lateralized regions (Hardwick et al., 2018) and tool responses were also found lateralized in the left hemisphere as reported before by Bracci et al. (2012). Verbs, action performance and tools all did not engage right LOTC for RFX GLM contrasts as expected.

The overlapping regions might be due to low spatial resolution of MRI resulting in responses at the same location, even though modules or neurons responsive to single domains are separate at a higher resolution. Future studies using higher spatial resolution could help to disentangle these. An additional explanation for the shared activation can also be that it is part of distributed but superimposed representations of all the different domains representing different features in the same location as Aflalo and Graziano (2006) as well as Graziano, Aflalo and Cook (2005) suggested. The used univariate analysis can mainly inform where the contrasts activate the brain, but it can tell us little about the representational content of this area to resolve this question. Therefore, multivariate analyses (Mur, Bandettini & Kriegeskorte, 2009) are valuable to address not only the shared location but also the shared representational content between domains in the LOTC. Consequently, multivariate pattern analysis (MVPA) was used to further investigate the organizational principles of LOTC and is reported in chapter 4.

Looking at the distribution of peak activations in all participants, we found a systematic spatial organization of most contrasts, with consistent locations of peaks for most participants for the single contrasts. Moving from anterior to the posterior edges of LOTC we found clusters of peak-responses for verbs, action observation and action performance. In the posterior LOTC we found overlapping clusters for body parts, tools, biological motion, basic motion and most posterior a cluster for peaks of the body contrast. Moreover, examining subregions of LOTC for selective responses to the different RFX GLM contrasts we found strong responses to only one or few

contrasts in only few subregions and instead for most of the central LOTC subregions up to six of the eight contrasts had a significant effect, and most regions showed the strongest response to action observation.

Furthermore, the investigation of the relation of the locations of responses to the different domains showed, that in line with a suggested gradient of abstraction from posterior to anterior regions (Watson et al., 2013, Lingnau & Downing, 2015, Wurm, Caramazza & Lingnau, 2017), the more concrete, visual domains like, motion, biological motion, bodies and tools were clustered at the posterior end of LOTC while more abstract concepts like action observation and verbs were found more anterior.

Looking not only at the location of the peak clusters but also at the distribution by focusing on the size of the confidence intervals and scattering of the peaks, some domains are more localized like verbs, action observation, body parts and tools with peaks clustering together across participants while basic motion and biological motion are widely distributed in occipitotemporal cortex. A more reliable location is more in line with a modular organization (Fodor, 1983; Fodor, 2001; Caramazza & Shelton, 1998; Kanwisher, 2010) as a similar region is involved in a similar process across participants while peaks spread across two lobes can be better explained by distributed processing (Haxby et al., 2001) as a number of different regions of interest show strong responses to the motion contrasts. In the posterior part a focal region seems to be located where most of the confidence intervals overlapped. The overlap speaks against clear cut borders between modules. These ambiguous findings could be integrated by a representational space with dimensions coding the different domains (Haxby et al., 2011; Kriegeskorte, Mur & Bandettini, 2008) where some dimensions could cover smaller regions while other dimensions are distributed across broader areas as Lingnau and Downing (2015) suggested. The region could be organized by superordinate dimensions with higher levels of generalization across categories on a larger scale that would result in gradients across the LOTC area, superimposed on more

categorical dimensions showing up as overlapping but somewhat selective subregions (Op de Beeck, Billet & Ritchie, 2019).

Given that the analysis of peak activation was implemented using a rather liberal (i.e. broad) region of interest including the occipital and temporal lobe, the result that the peaks for action observation were found to be located inferior to the other domains in inferior temporal cortex or fusiform gyrus for most participants was somewhat unexpected as these regions were not yet reported to be related to action perception specifically. In the meta-analysis by Caspers et al. (2010) as well as the meta-analysis from neurosynth reported in the introduction, the fusiform gyrus is associated with action observation but the close by region in LOTC is far more pronounced with a bigger and stronger cluster. One possible explanation is that in our action stimulus set, almost all pictures include one or two faces and as Casper et al. (2010) reported fusiform clusters only for face or whole-body stimuli, but not for hand stimuli, this cluster might relate to the presence of faces. To further investigate this, it would be interesting to examine the effect of different body parts in the representation of observed action to distinguish between effect of faces and of action understanding.

We argued that the responses to different tested domains overlap in some areas of LOTC and that they are topographically mapped across occipitotemporal cortex clustering at similar locations across participants and hemispheres. These results should be interpreted with care though, as we infer neuronal response from the hemodynamic signal in relatively large voxels. A driving factor of strong responses for several domains could also stem from the shared blood supply from middle cerebral artery and as larger vessels have a stronger signal (Polimeni, Fischl, Greve & Wald, 2010). This might influence the results and facilitate overlap and clustering of responses in the LOTC.

The analysis of selectivity of subregions within the LOTC region also revealed that some subregions, especially at the outer border of the LOTC including LO1, PH, PHT, TE1p and STSvp show responses selective to single or few domains, mainly action observation, tools, verbs or bodies, while the regions in the centre like V4t, MT, MST and FST responded to up to six tested domains and thus exhibited no selectivity. This again speaks for modularity in some way as single regions respond to single domains but a distribution of activation not lower than MRI resolution but instead spread across several subregions of LOTC was found, supporting a distributed representation with rich information about several domains coded in the centre of LOTC.

For the analysis of selectivity, the contrast for action observation (action pictures>scrambled pictures) showed the highest t-scores in most regions compared to the other contrasts. A possible explanation for this unexpected finding might be due to the fact that the action pictures were contrasted with pictures consisting of scrambled pixels only and possibly these activate only primary visual but not higher association areas resulting in strong results for the test condition due to the contrast. However, as body parts were contrasted with scrambled images as well and yielded weaker and more clustered results though, a second explanation seems more likely. The strong effect might be due to the rich stimulus of an action picture, consisting of one or two people in a scene, sometimes even using tools and the stimuli imply motion. The stimuli are a combination of several of the other tested domains and possibly this enhances the response of LOTC regions as superimposed representations might add up to an overall stronger result.

We have claimed that most regions in LOTC show overlapping, non-selective responses to different domains, but one might argue that these findings are not related to the domains but rather other differences in tasks or stimuli used in the different functional localizers. For example, the response for action observation dominated in most regions and the significant activation covered extensive occipito-parietal areas. For each task, we used a paradigm as

similar as possible to the other tasks by using simple block designs with blocks of 15 seconds and a one back task where applicable. Nevertheless, features of stimuli or tasks unrelated to the domains might differ, contributing to the differences of the responses. The action observation stimuli for example are rich in information consisting of real-life action scenes with bodies, body parts, objects, tools, scenes, faces and more and the different features together enable us to understand the action while a picture of a hand in comparison is rather plain. This difference between the stimuli might relate to the complexity of the presented information and even though we did not actually control or manipulate this stimulus aspect across the different functional localizers, the results of strongest responses for the most complex picture stimulus is in line with an additive, superimposed feature space suggested by Op de Beeck, Haushofer and Kanwisher (2008). For example, if bodies, hands and tools are features within the same space and their presence increases the response in LOTC regions individually, a concurrent presentation like in the action observation stimuli could lead to an additive increase of the response from all three features and therefore a stronger signal. However, as we included a range of different stimuli and tasks it is important to note that some differences between the tasks and stimuli that we did not control or consider in the analysis could influence the amplitude of the obtained responses.

Overall, the univariate analyses showed a considerable overlap for the domains of motion, biological motion, tools, body parts, bodies, action observation, action performance and verbs in the LOTC area and the different domains were clustered at reliable locations with a posterior part of LOTC including highly overlapping clusters of motion, biological motion, bodies, tools, body parts and more anterior individual clusters for action observation, action performance and verbs. The strength of responses to the individual contrasts in subregions of LOTC also showed regional differences with central regions recruited by most contrasts and anterior regions selectively recruited by single contrasts. Neither a sole modular nor a distributed organization

of the LOTC can incorporate the reported results. Instead we argue for superimposed spaces representing different features shared by the domains like Op de Beeck, Haushofer & Kanwisher (2008) and Aflalo & Graziano (2006) suggested. Possible feature spaces could relate for example to abstraction of information (Roth & Zohary, 2015; Wurm et al., 2017), real-world size, animacy (Konkle and Caramazza, 2013) or body parts (Orlov et al., 2015; Tarhan & Konkle, 2019). In our dataset we included high level retinotopy to further investigate the gradient of abstraction and potential additional analyses can utilise more detailed differentiation of body parts to investigate these features in LOTC in the future.

4 Representational organization in LOTC

4.1 Introduction

After evaluating the location of responses to different domains in the univariate analyses, we also complemented the results using multivariate analyses to investigate the representational content in different regions in LOTC. As we found substantial overlap of responses in the posterior and central parts of LOTC, we want to further investigate whether these shared regions also contain shared representations or whether different representations are located at the same location in the cortex.

Multivariate pattern analysis using representational similarity analysis can help to better understand not only where a contrast shows high activation but also how the domains are represented across voxels (Kriegeskorte et al., 2008). We want to outline the neuronal representational space in LOTC by using RSA to create representational similarity matrices (RSMs) in subregions and across LOTC signifying what these regions represent and how similar domains and regions are. First, we can use the resulting representations to investigate the sensitivity of regions for a domain indicated by high correlation of odd and even runs for the same domain (Ritchie et al., 2017). If a region is sensitive to a domain, it should show a reliable pattern in response to that domain. For modular organization, we would expect sensitivity to one or few domains per regions, while a distributed superimposed feature mapping would be expected to show sensitivity for several tasks across several regions.

Second, we can use the similarity of representations across domains to evaluate how similarly domains (or conditions) are represented and to detect clusters of domains where domains with similar neuronal coding are expected to be closer to each other in a representational space as well (Haxby et al. 2014). By evaluating what these domains share, we can infer what features

might be relevant in this region. By looking into the similarity of these representational spaces across regions, we can also investigate changes of representations as we would expect for a hierarchical organization gradually changing representations for example a gradual loss of spatial information as Roth and Zohary (2015).

The summarized applications of RSA refer to a model-free and data driven explorative approach that enables us to better understand how similarly different domains are represented and how similar representations are across different regions, but the analysis can also be used in a model driven way. For this, model RSMs are created in line with predicted similarity of conditions generated by a computational model or a hypothesis about the structure of the underlying feature space. The resulting model RSMs can then be compared to neuronal RSMs to establish whether the structure of activation patterns across the cortex are related to hypothetical structures. With this approach, it can be tested how well theoretical models align with functional organization of the cortex (Kriegeskorte, 2008).

4.2 Analyses

4.2.1 *Sensitivity, discriminatory property and similarity for domains in LOTC subparcels*

Representational similarity analysis was used to investigate how reliable the pattern of activation is across runs (Ritchie, Bracci and Op de Beeck, 2017) for each of the contrasts for basic motion, biological motion, tools, body parts, bodies, action observation, action performance and verbs. To evaluate the similarity of different contrasts in the subregions of LOTC, RSA is also used to look at the correlation of neuronal pattern across contrasts. We correlated the response patterns of t-values across voxels from each contrast of interest calculated in the previous GLM analysis in Chapter 3. For each participant, t-values from the first level analysis for each contrast of interest (see table 2 for contrasts) were extracted from each

of the individually defined HCP_MMP1.0 subregions (Figure 27). For each subregion, the t-values from odd runs were then correlated with results from even runs, resulting in a representational similarity matrix showing the similarity of contrasts within certain areas (Figure 28). We used custom Matlab scripts and the CosmoMOPA toolbox (Oosterhof, Connolly & Haxby, 2016). The resulting correlation coefficients were then Fisher-z-transformed, averaged across participants and retransformed into correlation coefficients (r-values).

To evaluate the discriminatory property (whether the representations can be used to discriminate between the contrasts) of the different subregions for the different contrasts. The discrimination index was calculated following Chan, Kravitz, Truong, Arizpe, & Baker (2010) by subtracting the average between-contrast correlations from the within-contrast correlations for each contrast and all regions of interest. To test for significant discrimination, we used one sample t-tests comparing the within-contrast correlation (value on the diagonal of the RSM) with the between/contrast correlations (off-diagonal values of the RSM) for every contrast and all subparcels.

We also performed an agglomerative hierarchical clustering analysis of all contrasts in each region of interest using the unweighted average Euclidean distance to compute the distance between clusters. We used dendrograms to visualise the resulting clustering of contrasts.

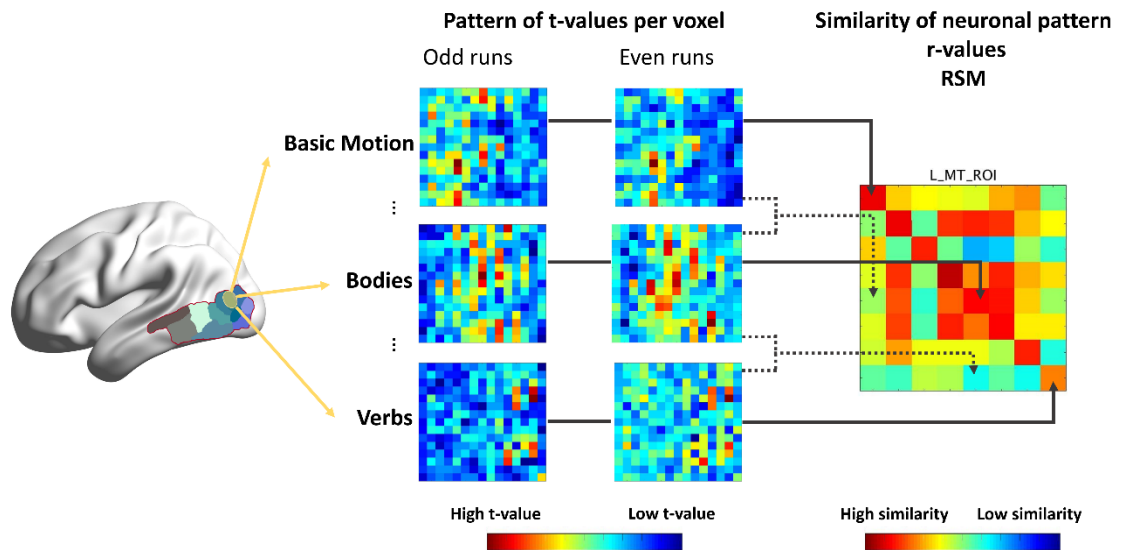


Figure 27. Representational Similarity Matrix (RSM) construction, here shown for region MT. The same procedure was used in each participant and region of interest. First t-values are extracted from the ROI voxelwise for odd and even runs and resulting patterns of t-values (every tile corresponds to a voxel) are correlated across runs and contrasts resulting in one RSM (every tile corresponds to a correlation r of response patterns).

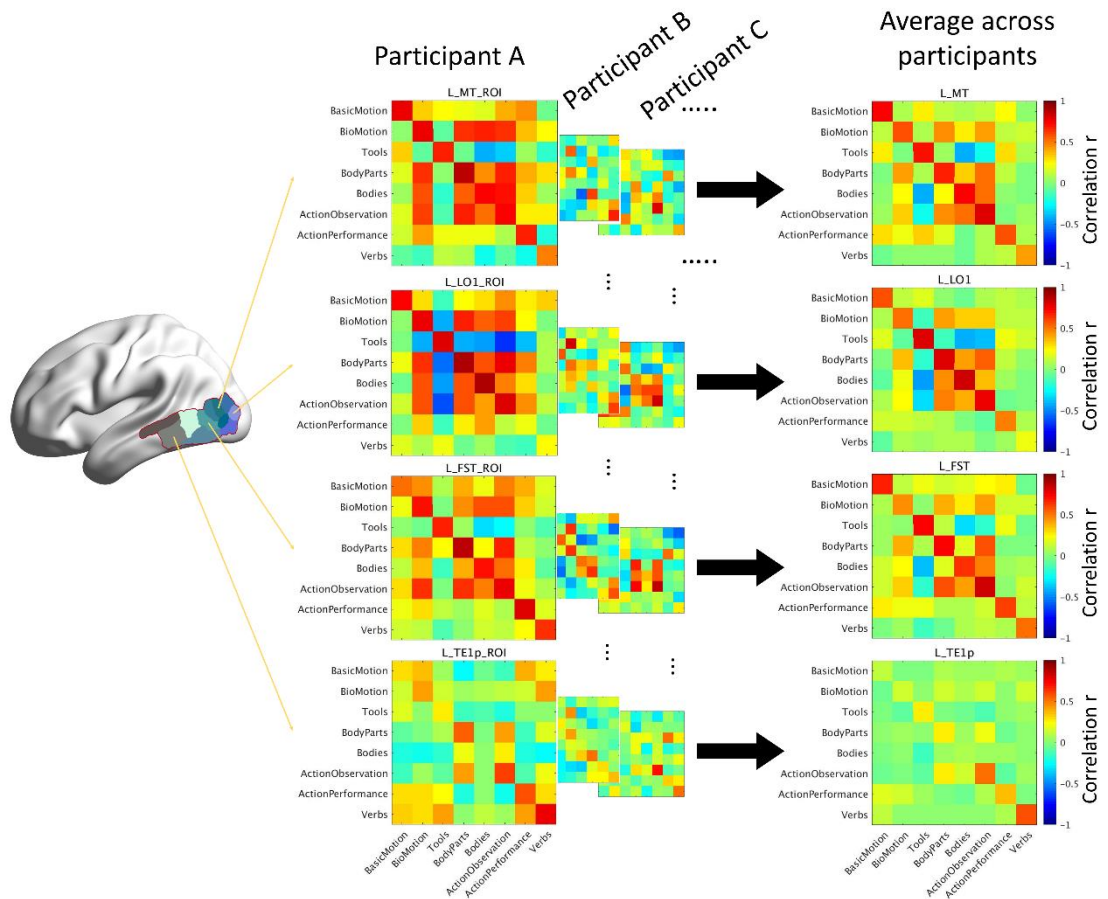


Figure 28. Average RSM construction. In every participant, t-scores were extracted from LOTC-ROIs for each run and each contrast of interest, odd and even runs were then correlated for participants individually within and across contrasts resulting in a representational similarity matrix (RSM). RSMs from each ROI were then averaged across participants.

4.2.2 *Similarity of representations across LOTC subparcels*

To assess the similarity of RSMs across regions a correlation of the RSMs from the previous analysis was calculated, comparing the representational similarity matrix from each region with all other regions within the LOTC. This can show similarity of different regions as well as changes of represented information across areas. Gradual changes of representation across the regions can be revealed like this or abrupt changes of decoded information. To avoid illusory effects driven by strong correlations in the diagonal of the representational similarity matrices (Ritchie, Bracci, Op de Beeck, 2017), we calculated the following steps separately for values in the diagonal and for values in the off-diagonal of the average RSMs in each region. Also the diagonal was used to investigate the sensitivity of single regions to the domains and the correlation across regions of RSM-diagonal values indicates how similar regions are in terms of what domains they represent. The off-diagonal values from RSMs however can be used to assess changes of representational similarity across domains allowing for a better understanding of possible underlying features changing across regions. RSMs from every ROI from the previous analysis were used and data from the diagonal and from the off-diagonal were correlated separately.

We then investigated the agglomerative hierarchical clustering of RSMs from different regions and calculated the distance between clusters using the unweighted average Euclidean distance and depicted this with dendrograms.

4.2.3 *HLR in LOTC-Subregions with representational similarity analysis*

The representational similarity analysis of the different domains in LOTC was then complemented with a representational analysis for the high level retinotopy task (see section 2.2.8 for details) we used to evaluate spatial sensitivity of LOTC. Specifically, with this analysis

we aimed to investigate whether information about stimulus position is represented not only in early visual cortex but also in LOTC areas.

For this, the preprocessing described in Chapter 2 was applied for the high level retinotopy task data and a general linear model-based analysis was performed per participant, session and run. The four locations (upper left, lower left, upper right and lower right) were modelled as block predictors convolved with a gamma function and 6 motion parameters from motion correction were included in the design matrix. Again, we used the HCP_MMP1.0 subregions and extracted t-values in every voxel in a given ROI for the contrasts for upper left vs. baseline (UL), lower left vs. baseline (LL), upper right vs. baseline (UR) and lower right vs. baseline (LR) for each run and participant. Representational similarity analysis was performed by correlating patterns of t-values from even and odd runs and results were Fisher-z-transformed, averaged across the group and retransformed into correlation coefficients. This was done for all 11 regions of interest.

Model representational similarity matrices were created for this experiment to evaluate whether, and in what regions, information about the position were represented. Models for the distinction between location in the left-right hemifield, upper-lower position or the single quadrant location were created (see Figure 29). For the model representing information about the hemifield location, responses for stimuli presented in the same half of the visual field are expected to be more similar across runs while responses for different hemifields show low similarity (Figure 29, panel A). If only the distinction between upper and lower half of the visual field is relevant for a given region, high similarity between the two upper conditions (UL & UR) is expected and respectively between the two lower conditions (LL & LR), but low similarity between upper and lower visual field combinations (Figure 29, panel B). Additionally, to test sensitivity to single quadrants in the visual field, the model assumes similarity between patterns obtained in odd and even runs when stimuli were presented in the same quadrant but low

similarity between different conditions (Figure 29, C). The neural representational similarity matrices were correlated with the model representational similarity matrices to evaluate how similar the models are to the neural response that we measured.

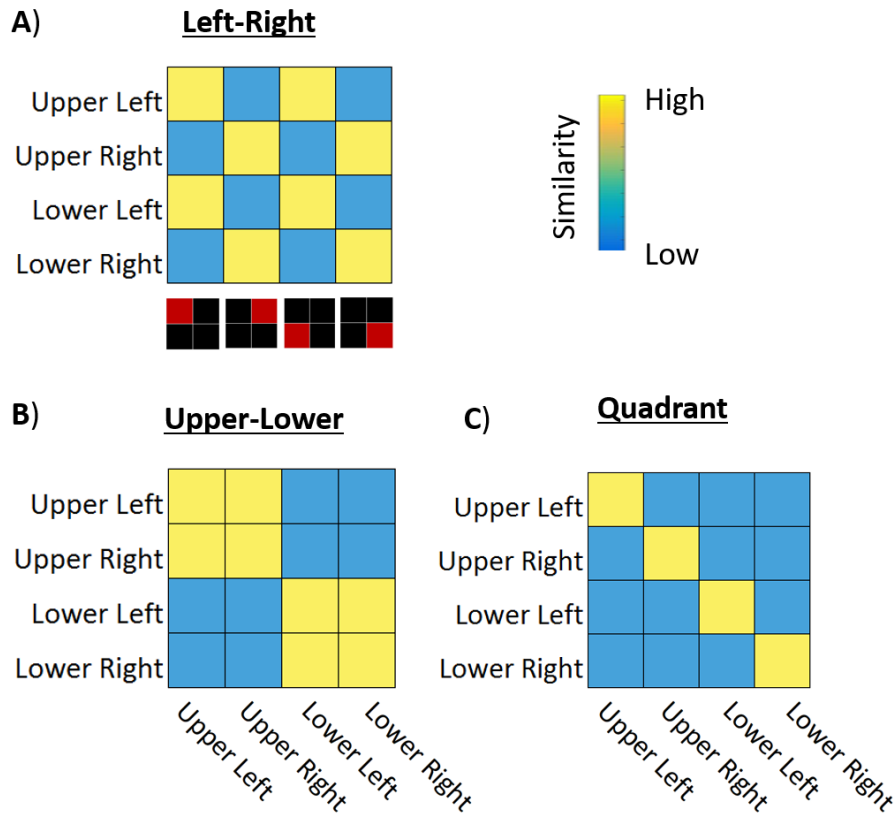


Figure 29. Schematic depiction of model-RSMs for patterns reflecting a distinction of stimulus presentation A) in the left or right hemifield, B) in the upper or lower half of the visual field or C) for distinction of each quadrant.

4.3 Results

4.3.1 Sensitivity, discriminatory property and similarity for contrasts in LOTC subparcels

To establish the within contrast reliability of neural patterns and how similar different contrasts are represented in one region, representational similarity analysis was used. High positive correlations in the diagonal indicate that even and odd runs show a similar pattern and therefore the neuronal representation is reliable across runs while no significant correlation shows that neuronal response during the task was different across runs and the activation is probably not related to the experiment but random noise (results are shown in Figure 30 as dot plots and Figure 32 and 35 as part of RSMs). Additional information about the discriminatory property of the different regions was assessed with the discrimination index where a high index shows a big difference in the representation (Figure 31, 33 and 36, and Table 6 and 7).

The posterior areas LO1, LO2, LO3, V4t, MT and MST areas show medium to high correlations of voxel-wise patterns of activity for most of the contrasts, while for the verb contrast LO1, LO2 and V4t exhibit low correlations (Figure 30). The verb contrast mainly shows high correlation across runs in the more anterior regions PHT, TE1p and STSvp of the left hemisphere (see Figures 30 & 32), but right hemisphere regions only show low to medium correlations signalling these regions do not reliably code this information (Figure 35). Verbs together with action observation and action performance are the only contrasts showing a reliable pattern of activity in left TE1p and STSvp. Only action observation and performance had neuronal patterns from odd and even runs correlating in the right hemisphere TE1p and STSvp while the other contrasts only generate low correlations between runs.

Motion can be significantly ($p < 0.001$) discriminated from the other contrasts in all (left and right hemisphere) but one region: STSvp in the left hemisphere (see Figure 33 and 36 and Table 6 and 7 for more detail). For biological motion, the representations for the contrast can be

discriminated from the other contrasts significantly in all regions. The discrimination index for the different contrasts (Figures 33 and 36) indicate for LO1, LO2, LO3, V4t, MT, MST, FST and PH a high discrimination for tools and decreasing indices from PHT, TE1p to STSvp. The representation of tools can be discriminated significantly from the other contrast in all regions (Figure 33 and 36 and Tables 6 and 7). For body parts a significant discrimination was found for all subparcels except right STSvp, for bodies all but left TE1p, right TE1p and right STSvp. The within-contrast similarity of action observation and action performance representations are significantly higher than the average between-contrast similarity in all regions of interest in both hemispheres. For verbs discrimination is significant in all left hemisphere subparcels and in the right in hemisphere in all regions except for LO1, LO2 and V4t (see Figures 33 and 36 and Tables 6 and 7 for more detail).

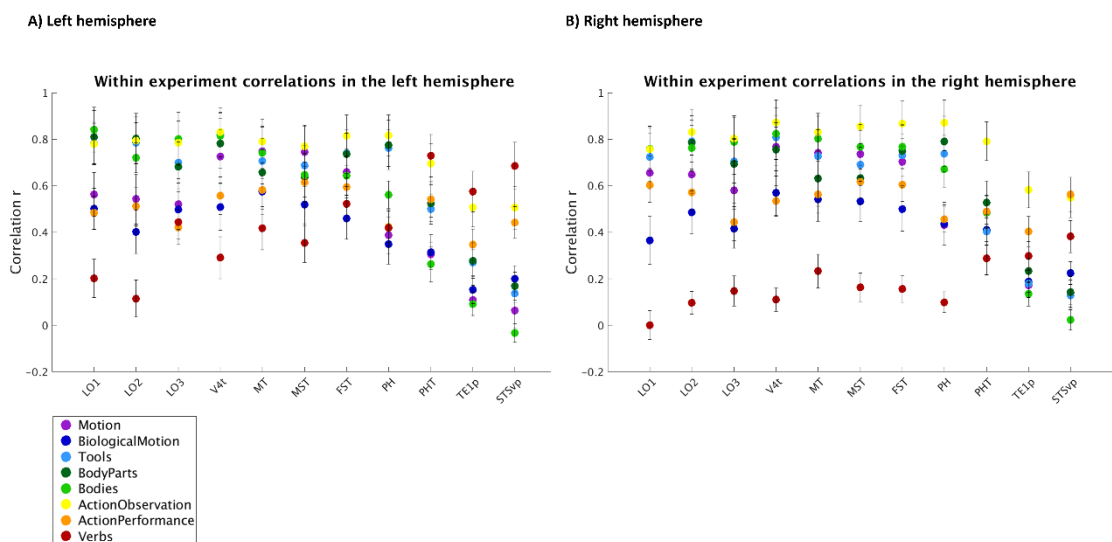


Figure 30. Sensitivity for contrasts in region depicted by correlation of odd and even runs for each contrast of interest across ROIs.

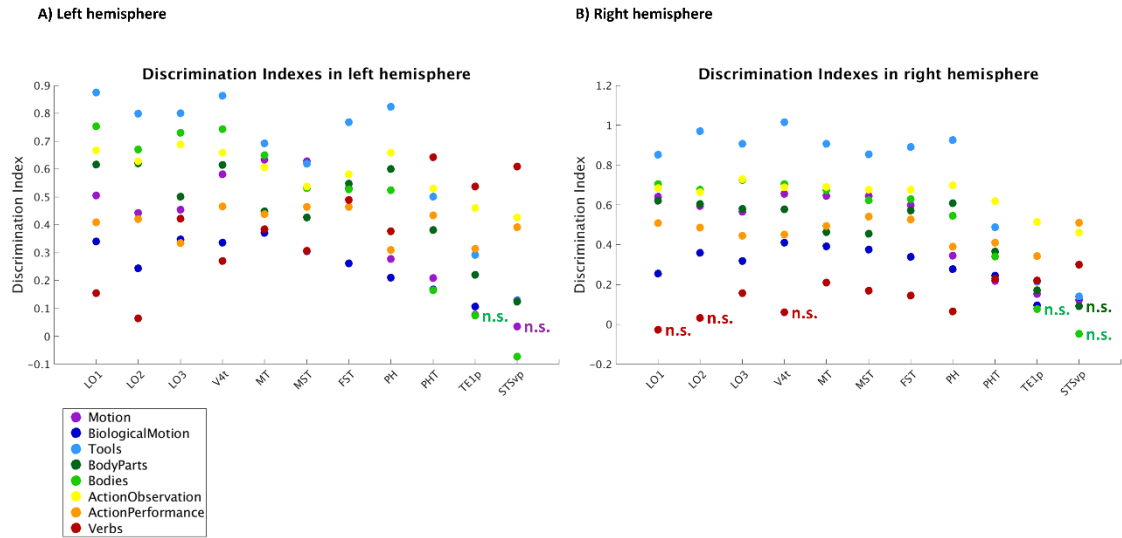


Figure 31. Discrimination index (within contrast correlation- between contrast correlation) for each contrast of interest across ROIs. n.s.= not significant difference between within and between correlations.

As can be seen in Figures 32 and 35, the patterns evoked by the observation of basic motion showed little similarity to the patterns evoked by any of the other functional localizers in both hemispheres. By contrast, biological motion was similar to body parts, action observation and a little less to bodies across the posterior and central regions in the left hemisphere. These four contrasts (biological motion, body parts, bodies and action observation) exhibited consistent and more pronounced similarity particularly in the central regions of the right hemisphere. Only for tools, we found negative correlations, indicating opposing activation patterns of the tool contrast and bodies and action observation contrasts, and low negative correlations for the pattern of activity for tools and biological motion and body parts. This was particularly distinct in posterior and central areas LO1, LO3, V4t, MT, FST and PH for both hemispheres. The contrast for body parts had an activation pattern similar particularly to action observation and a little less to biological motion and bodies in all bilateral subregions except the most anterior TE1p and STSvp. Bodies respectively activated posterior and central regions like body parts and action

observation. Action performance and verbs exhibited little to no shared representations and had marginal correlations between patterns of other contrasts in both hemispheres. Both hemispheres exhibited similar RSMs with the right hemisphere revealing stronger correlations overall.

Affirming the previous results, hierarchical cluster analysis revealed a main separation of two clusters consistent across LO1, LO2, LO3, V4t and MT (see Figure 34) in the left half of the brain and LO3, MT, MST, FST, PHT and TE1p in the right hemisphere (Figure 35). One cluster includes representations of biological motion, bodies, body parts and action observation showing a similar neuronal representation (also in MST, FST and PHT) and a second cluster with basic motion, action performance, verbs and tools. Recalling the RSM results, the first cluster has higher intertask-correlations, meaning the neuronal patterns of biological motion, bodies, body parts and action observation are similar to each other in the anterior regions but not similar to the basic motion, action performance and verbs contrasts, which show no systematic correlations to other contrasts. Patterns of activation for tools show the highest negative correlation with patterns for bodies and action observation and are part of the second cluster as well but patterns of t-values are less correlated with patterns from the other contrasts and in FST and PH (left hemisphere) and LO1, LO2 V4t, PH (right hemisphere) even make up their own, third cluster. In the anterior regions, TE1p and STSvp on the left and for the right only in STSvp the clusters break down as most of the correlation pairs show very little similarity between all contrasts resulting in high correlations between the homogenous RSMs.

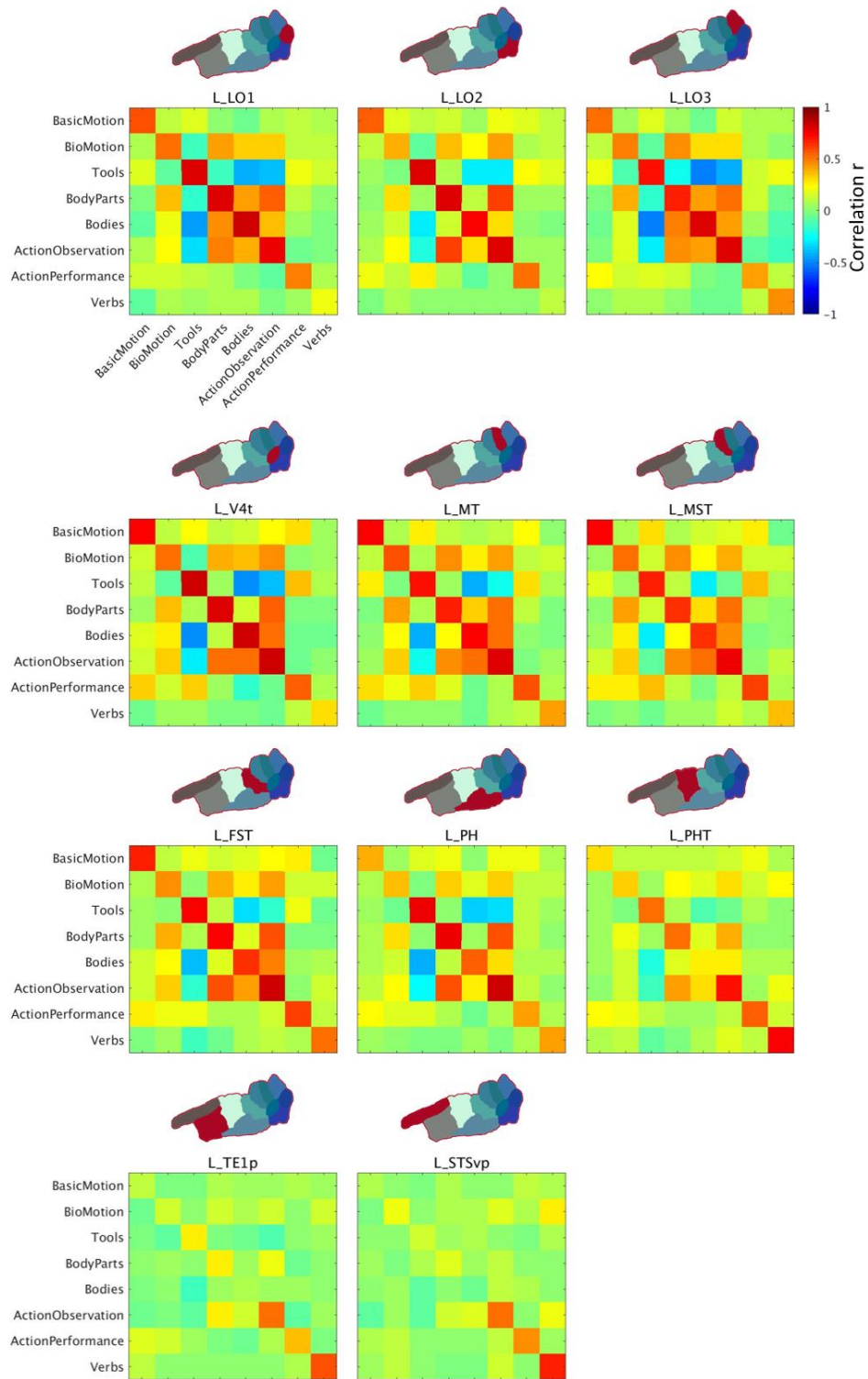


Figure 32. Average representational similarity matrices across participants in each region of interest in the left hemisphere. Single cells represent average correlation of distributed response patterns of experimental contrasts from odd and even runs depicting sensitivity and similarity of representations of domains in ROIs.

Table 6. Results for one-sample t-test to evaluate discrimination index (uncorrected for multiple comparisons) of all contrasts in the regions of interest in the left hemisphere. Correlation on- and off-diagonal values were extracted for each region of interest (ROI) from the representational similarity matrixes.

ROI	Contrast															
	Motion		Biological Motion		Tools		Body Parts		Bodies		Action Observation		Action Performance		Verbs	
	$t(13)$	p	$t(13)$	p	$t(13)$	p	$t(13)$	p	$t(13)$	p	$t(13)$	p	$t(13)$	p	$t(13)$	p
L_LO1	-22.727	<10 ⁻⁶	-7.630	<10 ⁻⁵	-13.507	<10 ⁻⁶	-9.427	<10 ⁻⁶	-9.699	<10 ⁻⁶	-8.704	<10 ⁻⁵	-19.769	<10 ⁻⁶	-10.024	<10 ⁻⁶
L_LO2	-25.353	<10 ⁻⁶	-6.546	<10 ⁻⁴	-15.414	<10 ⁻⁶	-11.407	<10 ⁻⁶	-13.605	<10 ⁻⁵	-9.064	<10 ⁻⁵	-16.477	<10 ⁻⁶	-4.669	<0.001
L_LO3	-18.945	<10 ⁻⁶	-8.095	<10 ⁻⁵	-11.235	<10 ⁻⁶	-7.162	<10 ⁻⁵	-8.292	<10 ⁻⁵	-8.727	<10 ⁻⁵	-14.382	<10 ⁻⁶	-17.056	<10 ⁻⁶
L_V4t	-22.496	<10 ⁻⁶	-6.977	<10 ⁻⁴	-11.683	<10 ⁻⁶	-11.658	<10 ⁻⁶	-8.927	<10 ⁻⁵	-7.723	<10 ⁻⁵	-10.573	<10 ⁻⁶	-20.190	<10 ⁻⁶
L_MT	-20.706	<10 ⁻⁶	-8.991	<10 ⁻⁵	-10.223	<10 ⁻⁶	-8.854	<10 ⁻⁵	-8.479	<10 ⁻⁵	-8.629	<10 ⁻⁵	-13.335	<10 ⁻⁶	-26.463	<10 ⁻⁶
L_MST	-22.086	<10 ⁻⁶	-8.024	<10 ⁻⁵	-11.252	<10 ⁻⁶	-8.900	<10 ⁻⁵	-8.331	<10 ⁻⁵	-9.661	<10 ⁻⁶	-13.454	<10 ⁻⁶	-18.050	<10 ⁻⁶
L_FST	-19.857	<10 ⁻⁶	-6.736	<10 ⁻⁴	-15.204	<10 ⁻⁶	-9.565	<10 ⁻⁶	-8.302	<10 ⁻⁵	-8.749	<10 ⁻⁵	-18.871	<10 ⁻⁶	-22.581	<10 ⁻⁶
L_PH	-14.922	<10 ⁻⁶	-6.467	<10 ⁻⁴	-15.319	<10 ⁻⁶	-11.082	<10 ⁻⁶	-9.611	<10 ⁻⁶	-9.253	<10 ⁻⁶	-17.961	<10 ⁻⁶	-35.947	<10 ⁻⁶
L_PHT	-12.935	<10 ⁻⁶	-7.520	<10 ⁻⁵	-20.191	<10 ⁻⁶	-10.177	<10 ⁻⁶	-4.773	<0.001	-12.096	<10 ⁻⁶	-24.409	<10 ⁻⁶	-29.774	<10 ⁻⁶
L_TE1p	-4.959	<0.001	-5.462	<0.001	-17.354	<10 ⁻⁶	-9.730	<10 ⁻⁶	-4.108	0.001	-15.593	<10 ⁻⁶	-16.096	<10 ⁻⁶	-48.195	<10 ⁻⁶
L_STSvp	-2.707	0.018	-6.305	<10 ⁻⁴	-9.338	<10 ⁻⁶	-10.194	<10 ⁻⁶	4.575	<0.001	-16.865	<10 ⁻⁶	-32.167	<10 ⁻⁶	-27.115	<10 ⁻⁶

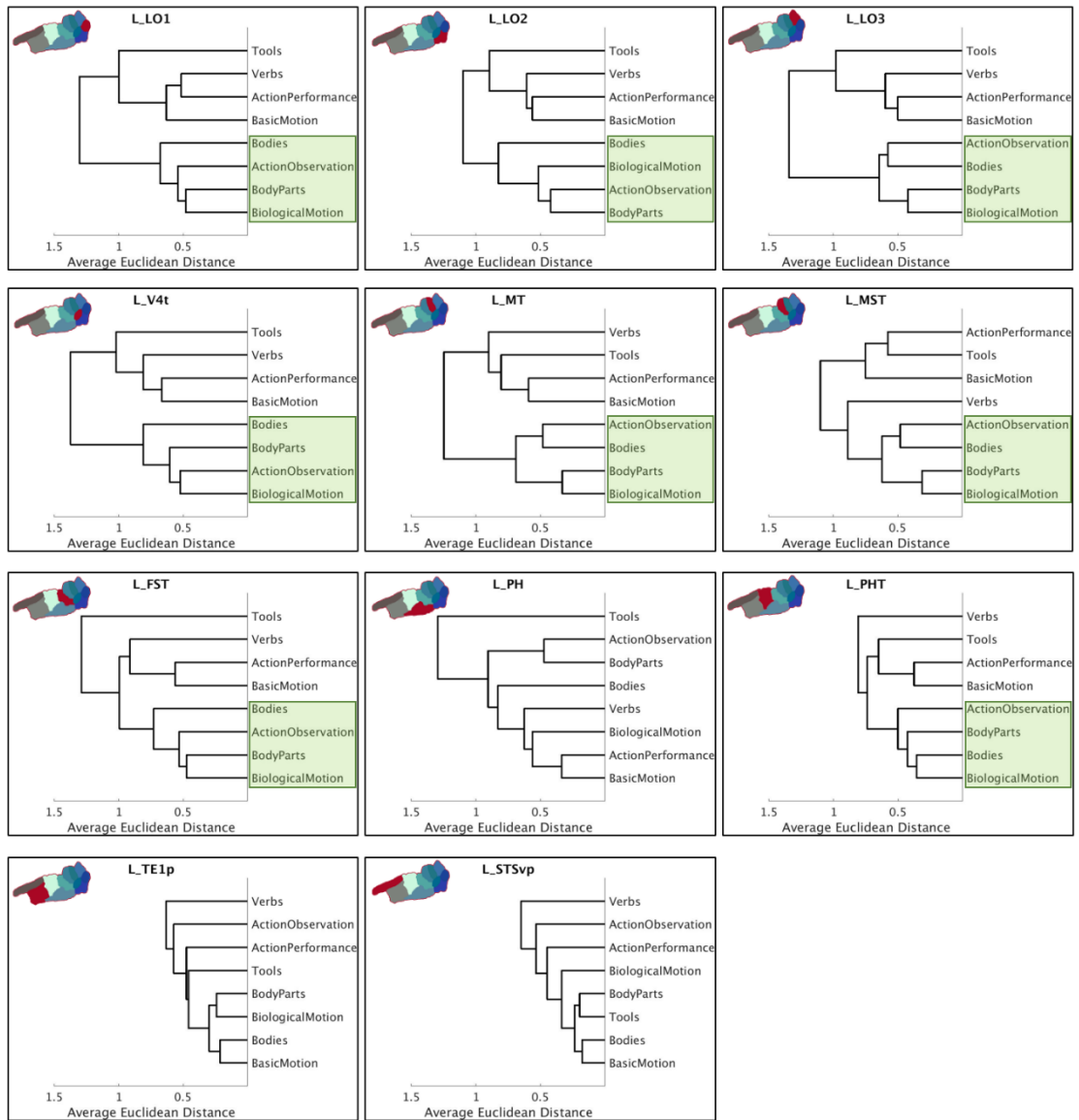


Figure 34. Dendrogram visualizations of the results from the left hemisphere based on hierarchical clustering of the correlations of distributed response patterns from experimental contrasts. Green box indicates recurring meaningful cluster of body related contrasts.

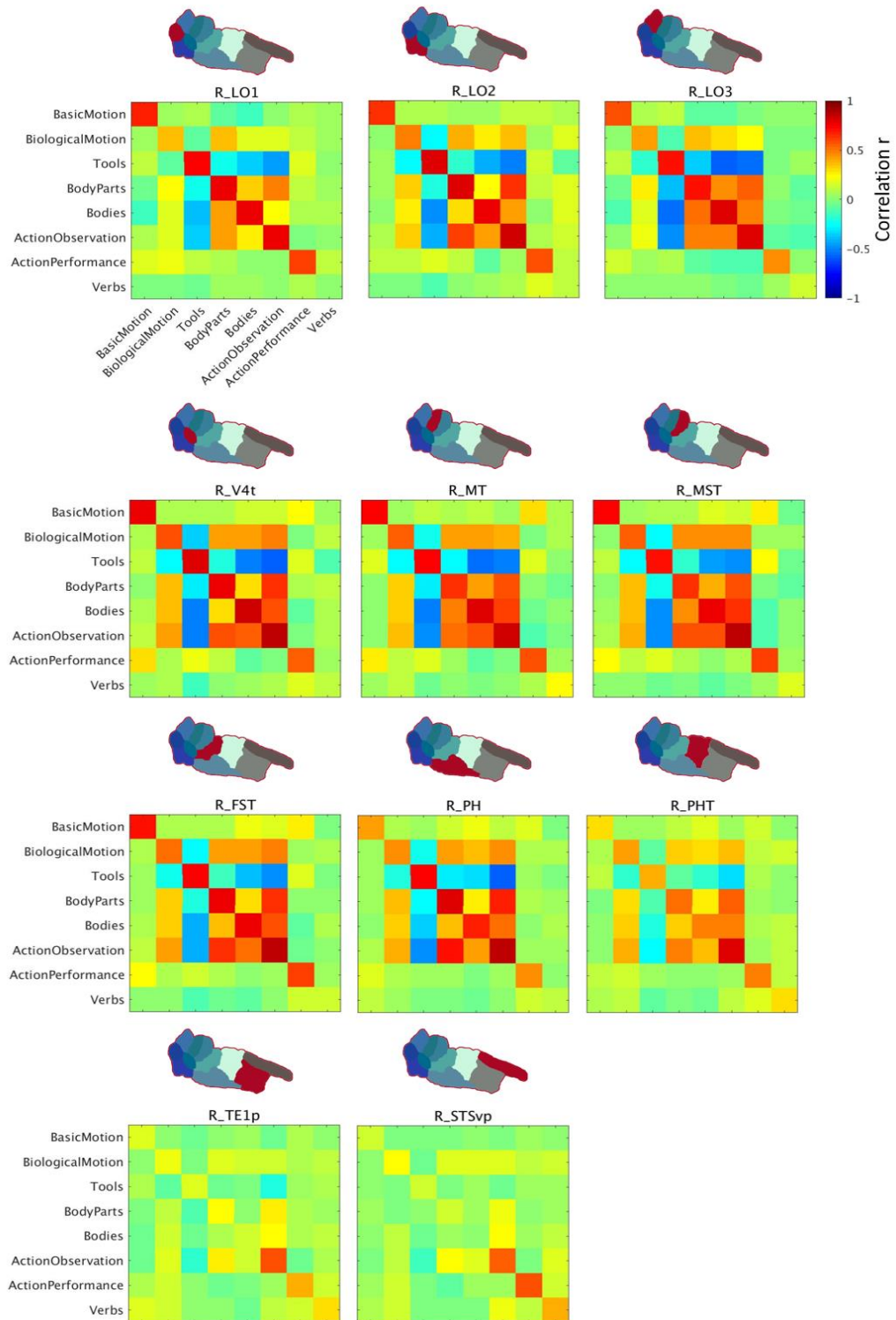


Figure 35. Average representational similarity matrices across participants in each region of interest in the right hemisphere. Single cells represent average correlation of distributed response patterns of experimental contrasts from odd and even runs depicting sensitivity and similarity of representations of domains in ROIs.

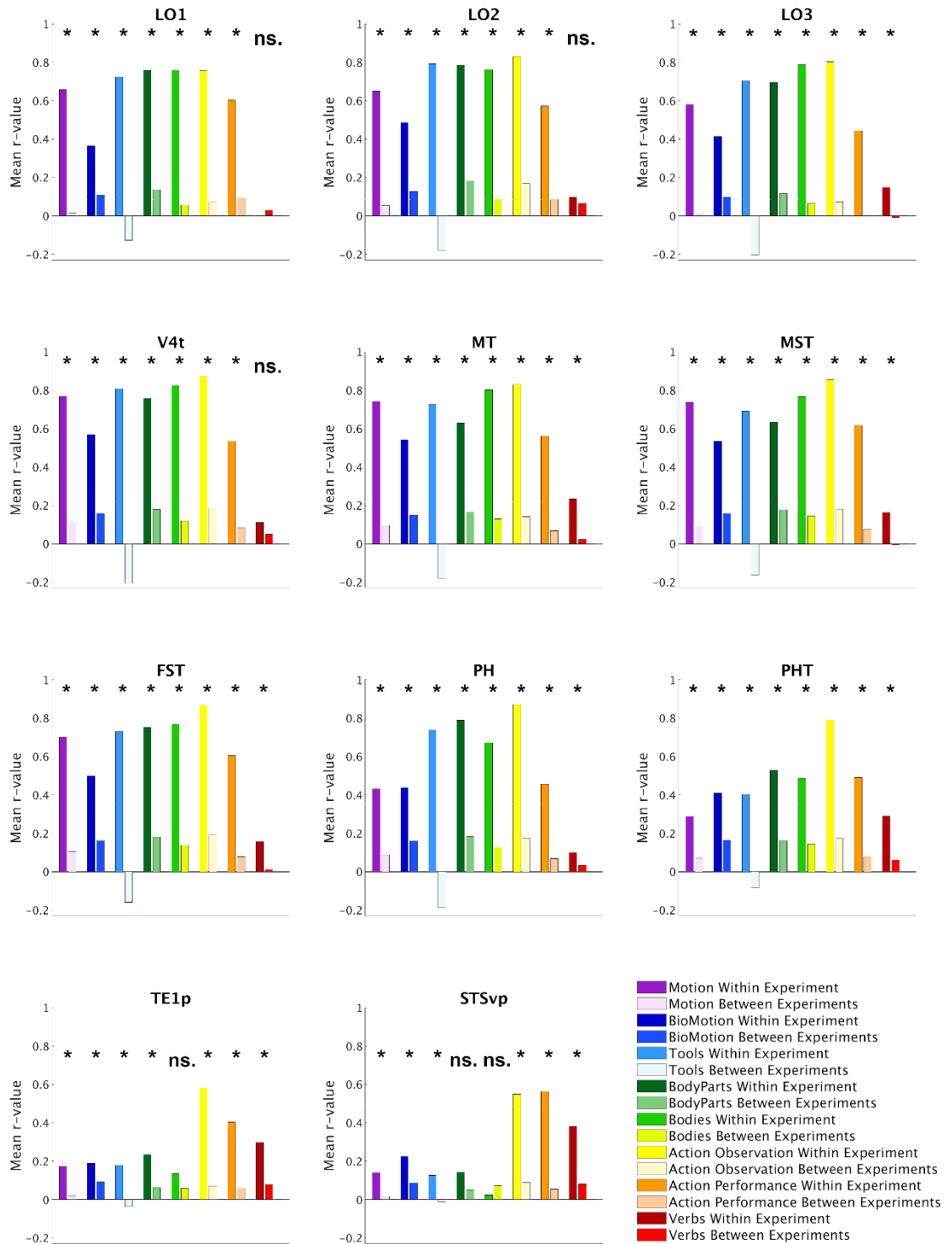


Figure 36. Difference between within contrast correlations and between contrast correlations for all regions of interest in the right hemisphere. *= $p < 0.001$; n.s.= not significant.

Table 7. Results for one-sample t-test to evaluate discrimination index (uncorrected for multiple comparisons) of all contrasts in the regions of interest in the right hemisphere. Correlation on- and off-diagonal values were extracted for each region of interest (ROI) from the representational similarity matrixes.

ROI	Contrast																				
	Motion			Biological Motion			Tools			BodyParts			Bodies			Action			Verbs		
	t(13)	p	t(13)	p	t(13)	p	t(13)	p	t(13)	p	t(13)	p	t(13)	p	t(13)	p	t(13)	p	t(13)	p	
R_LO1	-26.227	<10 ⁻⁶	-7.515	<10 ⁻⁵	-15.231	<10 ⁻⁶	-9.787	<10 ⁻⁶	-10.888	<10 ⁻⁶	-10.015	<10 ⁻⁶	-30.610	<10 ⁻⁶	2.413	0.031					
R_LO2	-64.016	<10 ⁻⁶	-6.327	<10 ⁻⁴	-15.451	<10 ⁻⁶	-8.910	<10 ⁻⁵	-9.095	<10 ⁻⁵	-7.317	<10 ⁻⁵	-26.620	<10 ⁻⁶	-1.304	0.215					
R_LO3	-28.042	<10 ⁻⁶	-7.464	<10 ⁻⁵	-12.752	<10 ⁻⁶	-6.940	<10 ⁻⁴	-7.652	<10 ⁻⁵	-7.991	<10 ⁻⁵	-22.170	<10 ⁻⁶	-13.427	<10 ⁻⁶					
R_V4t	-32.819	<10 ⁻⁶	-5.818	<10 ⁻⁴	-14.050	<10 ⁻⁶	-8.050	<10 ⁻⁵	-7.871	<10 ⁻⁵	-6.666	<10 ⁻⁴	-14.334	<10 ⁻⁶	-2.779	0.016					
R_MT	-27.422	<10 ⁻⁶	-6.448	<10 ⁻⁴	-12.333	<10 ⁻⁶	-6.022	<10 ⁻⁴	-6.838	<10 ⁻⁴	-6.908	<10 ⁻⁴	-15.267	<10 ⁻⁶	-17.019	<10 ⁻⁶					
R_MST	-27.776	<10 ⁻⁶	-5.529	<10 ⁻⁴	-12.591	<10 ⁻⁶	-6.030	<10 ⁻⁴	-6.589	<10 ⁻⁴	-6.672	<10 ⁻⁴	-17.300	<10 ⁻⁶	-16.222	<10 ⁻⁶					
R_FST	-27.867	<10 ⁻⁶	-5.208	<0.001	-15.040	<10 ⁻⁶	-7.969	<10 ⁻⁵	-7.691	<10 ⁻⁵	-6.814	<10 ⁻⁴	-20.508	<10 ⁻⁶	-8.920	<10 ⁻⁵					
R_PH	-22.695	<10 ⁻⁶	-4.547	<0.001	-16.246	<10 ⁻⁶	-7.493	<10 ⁻⁵	-7.614	<10 ⁻⁵	-6.835	<10 ⁻⁴	-26.454	<10 ⁻⁶	-4.275	<0.001					
R_PHT	-15.714	<10 ⁻⁶	-5.340	<0.001	-14.161	<10 ⁻⁶	-6.483	<10 ⁻⁴	-6.131	<10 ⁻⁴	-8.526	<10 ⁻⁵	-34.137	<10 ⁻⁶	-11.278	<10 ⁻⁶					
R_TE1p	-9.659	<10 ⁻⁶	-4.582	<0.001	-9.603	<10 ⁻⁶	-5.664	<10 ⁻⁴	-3.229	0.007	-12.600	<10 ⁻⁶	-20.264	<10 ⁻⁶	-12.729	<10 ⁻⁶					
R_STSvp	-15.164	<10 ⁻⁶	-6.648	<10 ⁻⁴	-9.505	<10 ⁻⁶	-3.849	0.002	2.443	0.030	-13.916	<10 ⁻⁶	-31.162	<10 ⁻⁶	-15.606	<10 ⁻⁶					

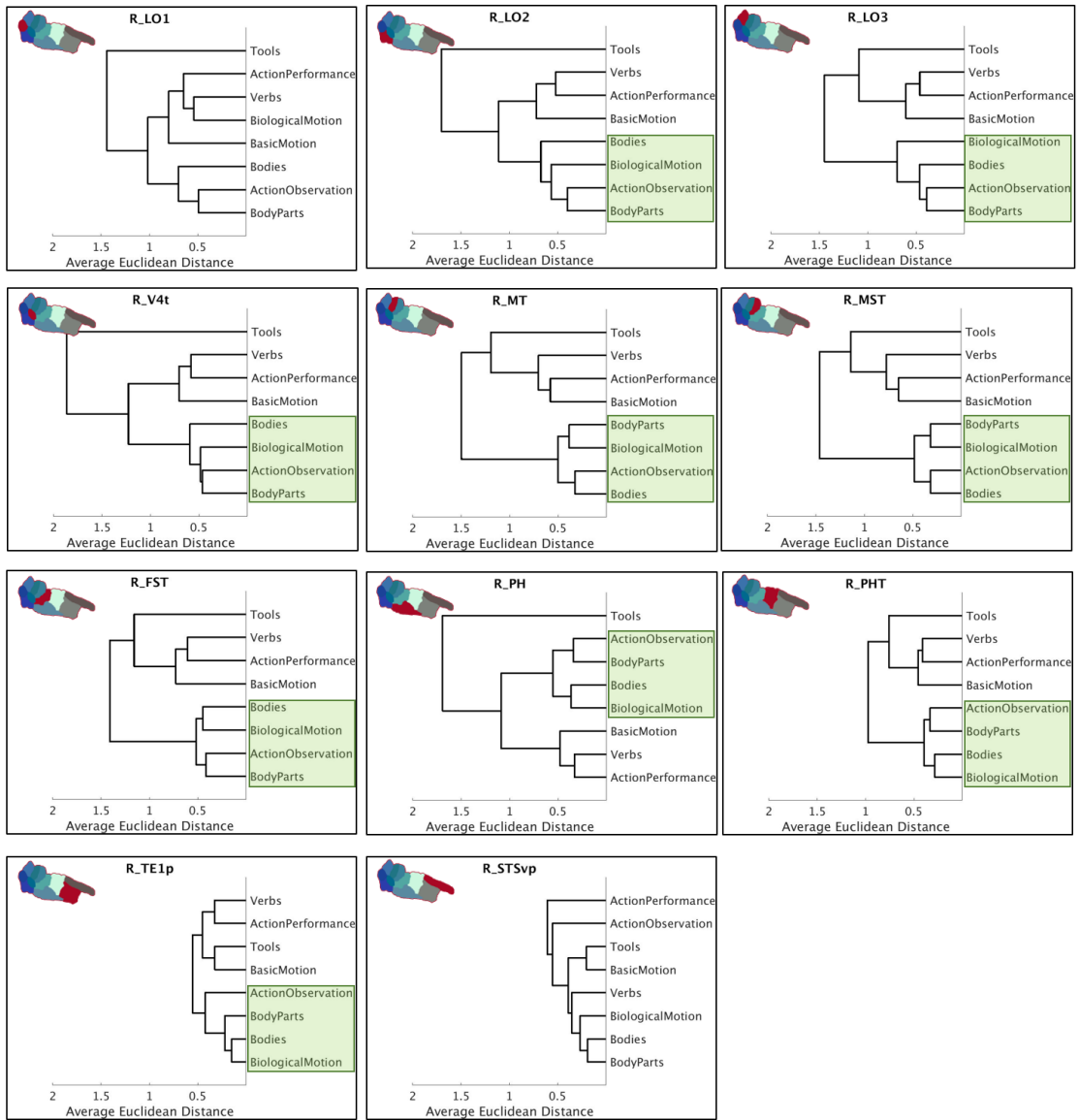


Figure 37. Dendrogram visualizations of the results from the right hemisphere based on hierarchical clustering of the correlations of distributed response patterns from experimental contrasts. Green box indicates recurring meaningful cluster of body related contrasts.

4.3.2 *Similarity of representations across LOTC subparcels*

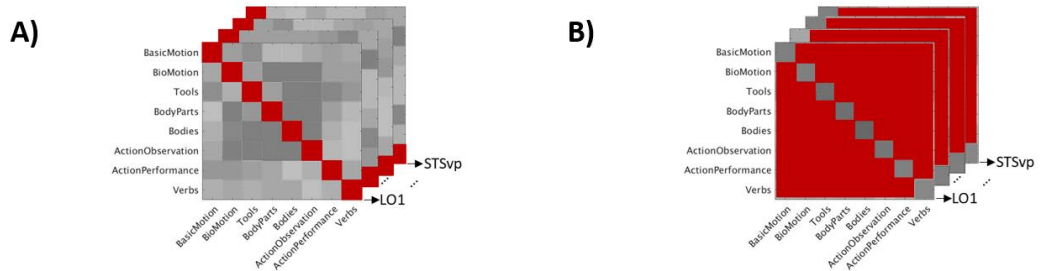
To evaluate the similarity of patterns obtained during the different functional localizers across LOTC subregions, we calculated the correlation of similarity matrices of the different regions and used dendrograms to visualize the distances.

Looking into the correlation and dendrogram (Figure 37, C & E) of the diagonal of the representational similarity matrices from each region, two clusters appear in both hemispheres. One cluster contains more posterior regions LO1, LO2, LO3, MT, MST, V4t, FST and PH (and PHT in the right hemisphere) while the second cluster includes PHT, TE1p and STSvp that are situated more anteriorly (only TE1p and STSvp in the right hemisphere). The posterior regions were sensitive to most contrasts while more anterior regions were sensitive to only verbs, action performance and action observation. Negative correlations were found between posterior and anterior regions as opposing sensitivity was found -for example no sensitivity for verbs posterior but high for bodies, while anterior regions show high sensitivity for verbs and no sensitivity for bodies in the left hemisphere of the brain. This was less prominent in the right hemisphere as the contrast for verbs was barely activating regions reliably.

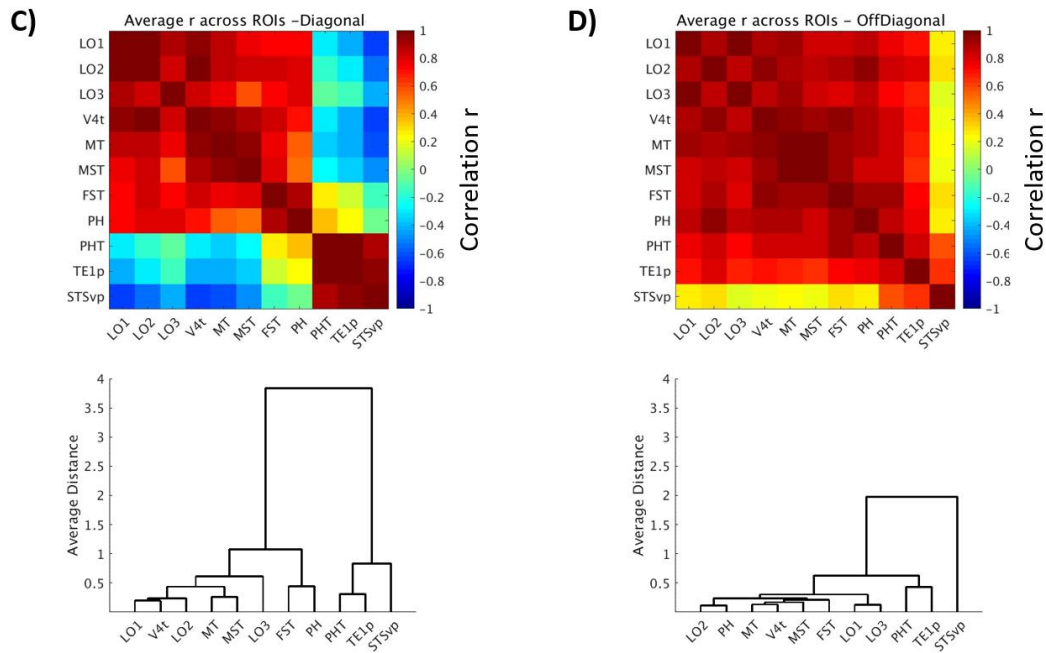
These two main clusters also showed subclusters with very high similarity in regions adjacent to each other like LO1, LO2 and V4t or MT and MST as well as FST and PH. Regions located close to each other seemed to be sensitive to similar contrasts but in the centre of LOTC between FST/PH and PHT the sensitivity changes abruptly even though the areas are very close.

The pattern offset from the diagonal in the representational similarity matrix shows high similarity across most regions of interest in the left hemisphere (LO1, LO2, LO3, V4t, MT, MST, FST, PH, PHT and TE1p) (Figure 37,D) but STSvp, the most anterior region, has patterns different from the other regions as can be seen in the dendrogram by the division in two main clusters. Similarity between different tasks is very alike in posterior regions and due to the fact that very

little cross-contrast similarity was found for anterior regions, they differ from posterior regions by hardly showing any correlations, especially STSvp. For the right hemisphere subregions of LOTC (Figure 37, F) anterior regions TE1p and STSvp clustered together and all central and posterior regions are part of a second cluster, but overall similarities across regions for the cross contrasts comparisons off the diagonal are high in the right hemisphere.



Left Hemisphere



Right Hemisphere

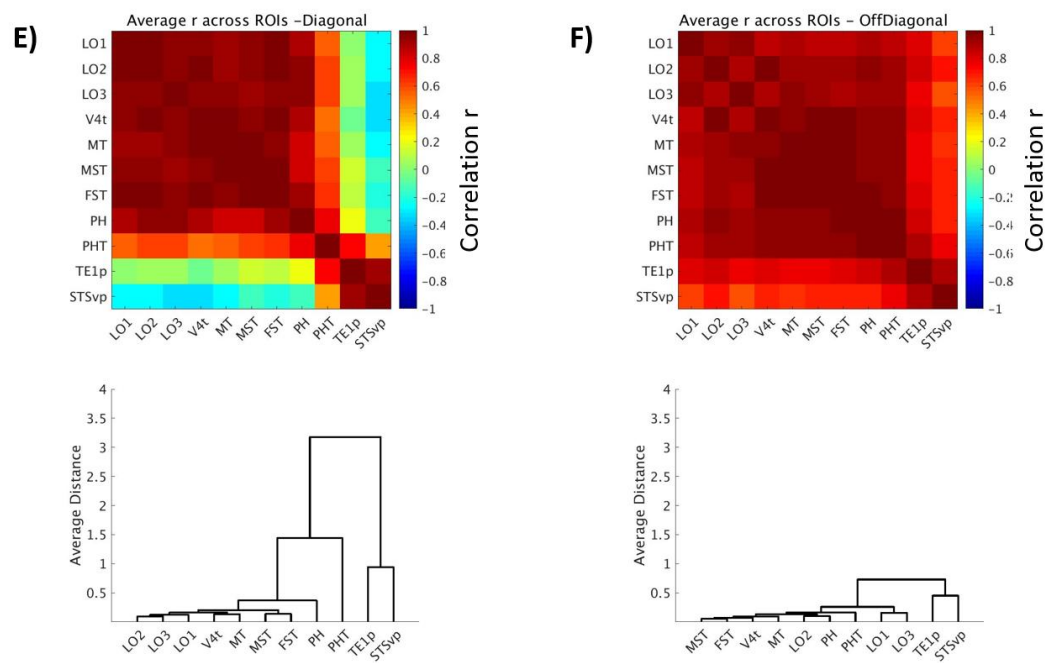


Figure 37. Similarity of RSMs across LOTC-regions for the domain contrasts. A) Schematic depiction of data used for analysis in red colour. B) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs. C) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs.

4.3.3 *HLR in LOTC-Subregions with representational similarity analysis*

The experiment to investigate the sensitivity to spatial position of complex stimuli, named “high level retinotopy” was subsequently analysed using representational similarity analysis as well. Different locations of stimulation (UL, LL, UR, LR) were used and pattern of activation in response to the different positions were correlated in the LOTC subregions in both hemispheres. The response pattern between even and odd runs indicated similar, reliable activation for the same conditions was highest in the posterior regions LO1 and V4t and continuously declined in LO2, LO3 as well as MT, MST, FST and PH and there is no sensitivity for the location in PHT, TE1p and STSvp in left (Figure 38, A) and right (Figure 38, B) hemisphere. In LO1, the similarity between different locations showed low correlations between conditions that shared either hemifield or height, for example like upper left and upper right and negative correlations were found for conditions not sharing hemifield or height at all like upper left and lower right. The more anterior the regions are located, the higher are the similarities in the diagonal and similarities between conditions with the same hemifield, while combinations of conditions from different hemifields exhibit negative correlations. In the anterior regions PHT, TE1p and STSvp no considerable similarities were found.

4.3.4 *HLR across regions*

A correlation followed by hierarchical clustering of RSMs from the different regions for the diagonal representing sensitivity for each location shows a solution with several small clusters with small distances (Figure 39, C & E). LO1 exhibited the strongest sensitivity for all conditions and makes up an own cluster in the LH, while it is similar to V4t in the RH. In the left hemisphere, a big cluster of regions (LO2, LO3, V4t, PH, PHT and TE1p) surrounding the cluster of MST and MT had lower sensitivity to the specific location of the stimulus. STSvp is represented by an own cluster. In the right hemisphere one big cluster is formed of posterior regions LO1, LO2, LO3,

V4t, and central region FST complemented with an anterior-superior cluster of MT, MST, PHT and STSvp while the two inferior regions TE1p and PH make up own clusters.

For the similarity across regions with data off the diagonal (Figure 39, D& F) – the similarities between the location conditions we find two strongly distinct clusters with all posterior and central regions in one and TE1p and STSvp in the second cluster in both hemispheres. In the first central-posterior cluster similar subclusters were found across hemispheres where LO1 makes up an own cluster as it showed negative correlations for conditions not sharing a dimension and low correlation for conditions sharing one dimension. This pattern changes as the hemifield dimension gets higher correlations if shared in the other posterior and central regions (LO2, LO3, V4t, PHT, MT, MST, FST, and PH).

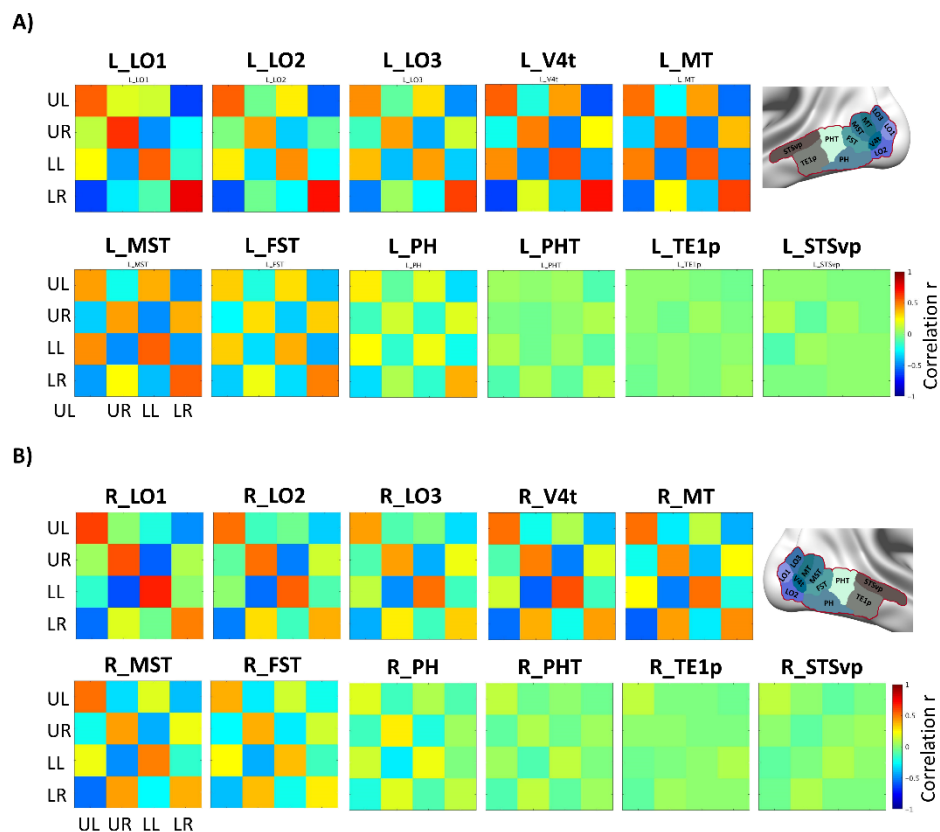


Figure 38. Average representational similarity matrices across participants in each region of interest in the right hemisphere for the spatial information experiment. Single cells represent average correlation of distributed response patterns of experimental conditions depicting sensitivity and similarity of representations for different locations of stimulus presentation.

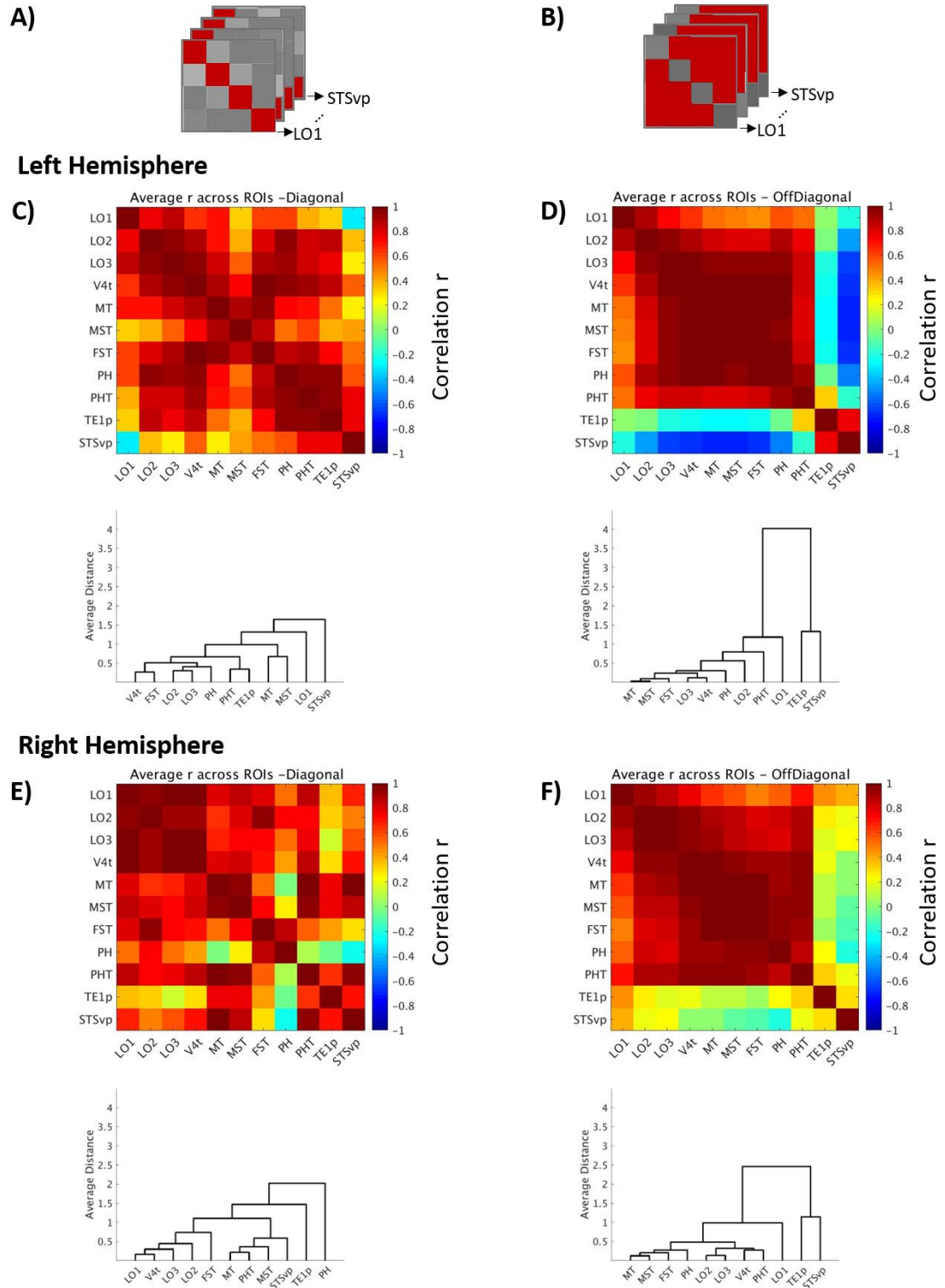


Figure 39. Similarity of RSMs across LOTC-regions for the spatial location contrasts. A) Schematic depiction of data used for analysis in red colour. B) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs. C) Average RSM and dendrogram visualisation for cells on and off the diagonal in left hemisphere ROIs.

4.3.5 *Representational content evaluated by models*

By testing three different models of location sensitivity – hemifield, upper-lower or quadrant models, we wanted to investigate what spatial information about the location of stimuli is coded even outside of early visual cortex in the LOTC. Two models covering half of the visual field each (left/right half or upper/lower half) and one model of higher resolution testing quadrants of the visual field were used.

The hemifield model revealed high similarity of neural patterns in central LOTC regions (MT, MST, V4t and FST) and the model RSM, surrounded by correlations a little less strong in LO1, LO2, LO3 and PH while we found a steep drop of similarity of neuronal response and model in PHT and close to no correlation for anterior regions TE1p and STSvp (Figure 40, panel A).

The second model evaluated sensitivity to upper versus lower visual field. A gradual decline was found from the most posterior regions LO1, revealing the strongest similarity of neuronal pattern and model for upper/lower half of the visual field through LO2, LO3, V4t, with medium correlations to MT, MST, FST, PH, PHT, TE1p and STSvp resulting in low correlation in both hemispheres (Figure 40, panel B).

Testing for higher spatial resolution the model of individual quadrants yielded high correlations for posterior and central LOTC regions (LO1, LO2, LO3, V4t, MT, MST, FST) with a decline in PH and PHT (in the right hemisphere) and very low similarity of model and neuronal response in anterior regions TE1p and STSvp (Figure 40, panel C).

Overall the highest correlations were found for the hemifield model in central LOTC, indicating that information about the hemifield in which a complex stimulus such as an action is seen is coded in LOTC and there is a rather sudden loss of this information in the anterior regions PHT, TE1p and STSvp.

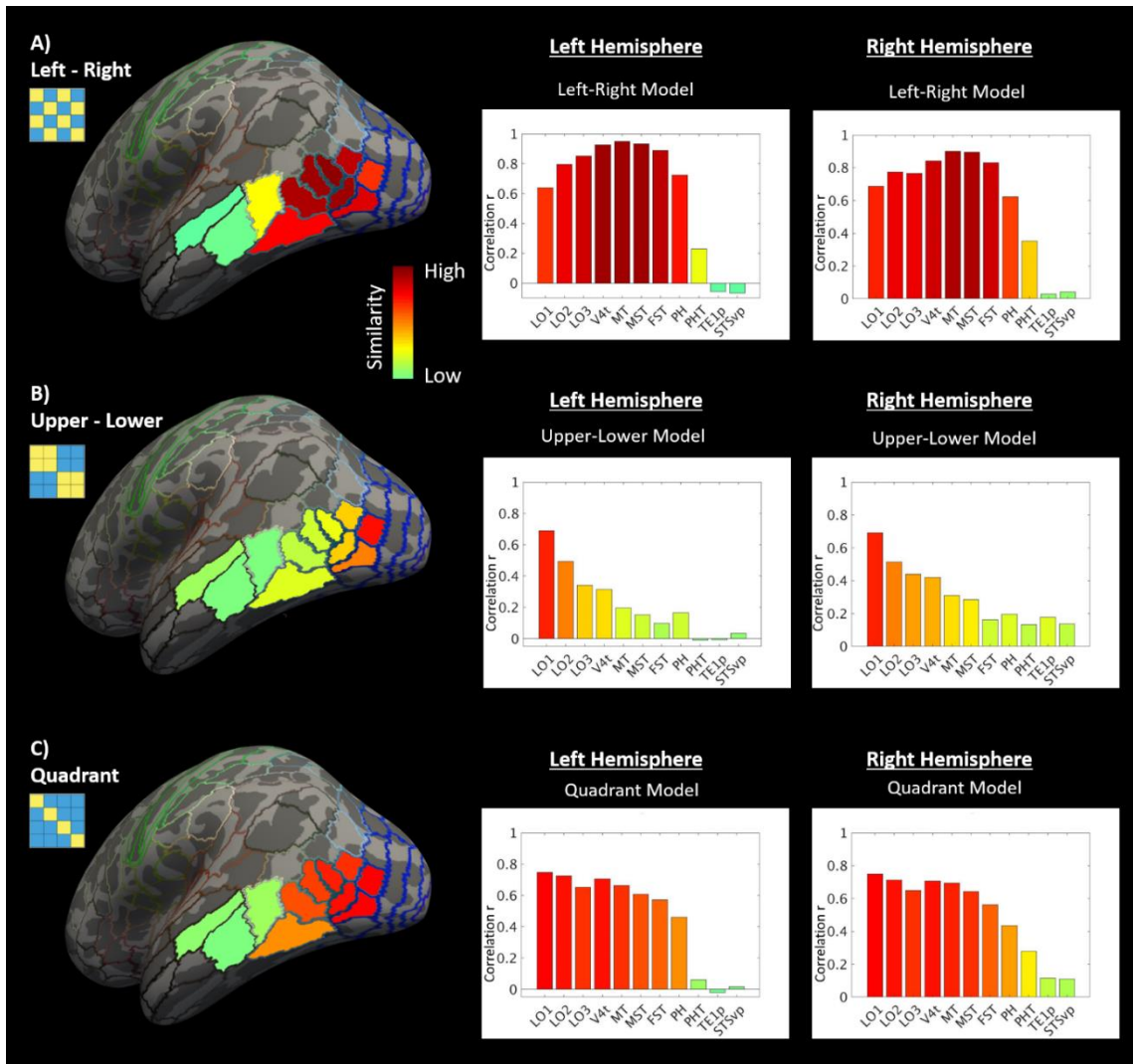


Figure 40. Model evaluation for response patterns to the spatial location experiment in LOTC-ROIs visualized on surface parcellation in the left hemisphere and bar chart for both hemispheres. A) Correlation of response pattern and hemifield-model. B) Correlation of response pattern and model of upper and lower visual field distinction. C) Correlation of response pattern and quadrant-model.

4.4 Discussion

The current analyses set out to explore the representational similarity of different tasks in and across LOTC subregions. This can provide an indication of (a) the domains to which subregions are sensitive and show discriminatory properties, (b) where different domains are represented similarly and (c) how the sensitivity and similarity change (gradually or abruptly) across regions in the LOTC. Our results indicate that the posterior and central subregions of bilateral LOTC contain rich information about basic motion, biological motion, tools, body parts, bodies, action observation, action performance and more centrally also verbs, while the anterior regions showed reliable information coding only for action observation, action performance and verbs (verbs were lateralized in the left hemisphere). This distinction could be interpreted as two modules, processing different kinds of information. Another possibility is that not only two kinds of information are processed, but this division separated different representational spaces where numerous kinds of information can be represented, but they differ in these two broader regions. This is also supported by the same boundaries found for the RSA of the high level retinotopy and the fact that information about the stimulus location was found in posterior-central regions but not in anterior ones. The representational space in the posterior-central LOTC includes not only features relating to the different domains, but also the location of stimuli in the visual field while the anterior regions seem to be involved in different processes for example related to semantic verb processing.

The discrimination index showed a significant distinction between the tested contrast and all other contrasts for most regions and contrasts. This means that the representation for visual motion can be distinguished from the average representation of all other contrasts combined and indicates that all contrasts show some unique pattern in nearly all LOTC subregions.

The similarity between different contrasts illustrated a shared representation of information for the body parts, bodies, biological motion and action observation while the other contrasts did not activate patterns very similar to any other contrast. This cluster of body-related stimuli is in line with initial reports about EBA, responding to bodies and body parts in all kinds of depiction (Downing et al, 2001). More recently, Hafri et al. (2017) reported generalisability of static and dynamic action depiction in LOTC and this fits with our finding that some information from moving point light-displays is similar to static depiction of bodies and whole actions.

Looking into the sensitivity of the regions, posterior and central regions (LO1, LO, LO3, Vt, MT, MST, FST, PH) show high similarity and seem to represent information similarly while anterior regions (TE1p, STSvp) are similar to each other as well but represent different information. Less pronounced but still visible is the difference between anterior STSvp and partially also TE1p and PHT for the similarity of representations across different domains.

As subregions of LOTC show very similar sensitivity to different kinds of information in the posterior and central parts, our findings in these regions are in line with a gradual, distributed representation of the different domains (Lingnau & Downing, 2015) and as several regions show similar representation of the body-related domains one important dimension of the representational space could be the involvement of human body (in also in motion or only parts of it). As the same areas were also sensitive to tools, motion and action performance contrasts a superimposed representation of different features as proposed by Aflalo & Graziano (2006) as well as Graziano, Aflalo and Cook (2005) can incorporate these different findings.

We also investigated the spatial location sensitivity of LOTC with a high level retinotopy to evaluate whether and what information about stimulus location is retained in LOTC. This again revealed sensitivity of posterior and central regions of LOTC to the location of a stimulus in the visual field while anterior regions do not code this information reliably. The similarity across the

location conditions indicated that the conditions with the same hemifield were represented similarly (UL and LL/ UR and LR). This was supported by the analysis of three different model RSMs, differentiating between information about the position in the dimension left-right (hemifield), upper-lower half of the visual field or quadrant specific information. The neural response pattern best matched with the hemifield model resulting in high correlations in posterior and central LOTC subregions.

The drastic change of representations from posterior and central to anterior regions was also reinforced by the comparison of the different subregions as we found an abrupt change of similarity again for the high level retinotopy experiment focusing on the cross-condition patterns. TE1p and STSvp did not convey reliable information about stimulus location and no similarity of conditions was consequently found. A previous study by Roth and Zohary (2015) investigating the spatial location sensitivity in higher visual areas found results indicating a gradual loss of information about the location from posterior to anterior regions in occipito-temporal cortex and while we also find that information about the location is coded in LOTC areas and the representation changes gradually, more surprising was the abrupt loss of information at the regions PH, PHT to just about no information in TE1p and STSvp. This finding supports a gradient within posterior and central regions of LOTC but a clear break between central and anterior regions is present and indicates a border of a representational space in LOTC.

Only the within-condition similarity across regions indicating the sensitivity for spatial location did not exhibit a clear distinction of anterior and posterior/central regions but this might be due to the sparse differences between conditions and the small number of conditions. Single regions showed overall highly similar sensitivity to each of the conditions and even though these correlations got weaker across regions – the pattern across four similar values does not seem to be enough to indicate meaningful similarity or dissimilarity across the regions.

Overall, we also did not only find indication of a rather sharp boundary between anterior and central regions of LOTC which would be in line with a modular organization, but we found gradual shifts of sensitivity and similarity across domains across the subregions in the posterior and central regions. The sensitivity for some tasks like verbs increased steadily from posterior to anterior regions and biological motion showed a gradual decline in regions more anterior than MT. Additionally, regions spatially close also tended to cluster together based on the neuronal pattern similarity they exhibit –posterior LO regions and V4t often clustered with each other while more central MT, MST and FST often clustered with each other and also V4t – bridging posterior and central regions and FST also clustering with the more anterior PH region and for some analyses with PHT. This speaks for representations being not localized in one region but spreading over numerous subregions leading to similar but gradually changing response patterns in the posterior and central part of LOTC, while anterior regions are not part of this representational space as they exhibit distinct sensitivity and pattern similarity.

5 Connectivity based organization in LOTC

5.1 Introduction

In this chapter, we want to complement our previous investigation of the location and the pattern of task-based responses in LOTC with the analysis of the task-free functional connectivity of LOTC to better understand the organizational principle underlying neural responses of LOTC. Resting-state functional connectivity uses spontaneous activation patterns of the brain, implemented as time courses of BOLD-signal, to find similarly activated brain regions. To examine the functional connectivity of regions or voxels, correlation of the responses of regions or voxels with each other are used and similar time courses indicate a functional connection of regions (Van Den Heuvel & Pol, 2010).

As we have reviewed in the introduction, for processes like motion, tool and body perception or action performance, action and verb understanding, not one single brain region is involved for processing these domains, but networks are involved in processing these domains. In the literature we found LOTC to be part of respective networks of all our domains - for visual motion (Culham et al., 2001), biological motion (Grossman et al., 2010; Saygin, 2007; Saygin et al., 2004), tools (Peelen et al., 2013; Reynaud et al., 2016; Lewis, 2006; Watson & Buxbaum, 2015), Bodies (Ramsey, 2018; Hodzic et al., 2009; Amoruso et al., 2011), action observation (Caspers et al., 2010; Grezes & Decety, 2001; Hardwick et al., 2018), action performance (Hardwick et al., 2018, Gerardin et al. 2000; Orlov et al., 2010; Gallivan et al., 2011) and language (Turkeltaub et al., 2002; Mechelli et al., 2005; Fedorenko & Thompson-Schill, 2014; Hagoort, 2014).

Previous research by Turken and Dronkers (2011) on the connectivity of regions that are part of the LOTC, such as the middle temporal gyrus, reported a gradual change of connected regions from posterior regions, strongly connected to occipital and parietal regions but quite narrow

spread, to anterior regions exhibiting more extensive connections in frontal, parietal and temporal cortex.

As these networks do share one node in LOTC, but differ in terms of what other regions are part of the network (see table 1 for an overview), we want to investigate to which regions the subregions of the LOTC are connected and how similar these connections are across regions. This enables us to investigate whether the functional task-based parcellation of LOTC into an anterior and a posterior-central division aligns with the functional connectivity patterns and possible divisions based on it in the LOTC.

5.2 Resting state data analysis

5.2.1 Preprocessing

Resting state data were preprocessed using FMRIB's Software Library (FSL; www.fmrib.ox.ac.uk/fsl), and involved the following steps: motion correction (MCFLIRT – Jenkinson et al., 2002), masking of non-brain voxels (BET; Smith, 2002), smoothing with full-width half-maximum 6 mm; high-pass temporal filtering with a cut-off of 100s and custom slice time correction.

Clearing data of motion artifacts using independent component analysis (ICA) has been shown to improve data quality (Birn et al., 2008, Pruim et al., 2015, Starck et al., 2013). To identify and remove noise signals, we therefore used ICA-based automatic removal of motion artifacts (ICA-AROMA) (Pruim et al., 2015). Independent components were extracted by FSL's Multivariate Exploratory Linear Decomposition into Independent Components (MELODIC) tool and single components for each subject were classified by ICA-AROMA as either signal or motion related noise (Pruim et al., 2015). Motion components were classified based on four features: high frequency content, maximum correlation to realignment parameters, edge fraction and cerebrospinal fluid fraction.

While blood oxygenation level-dependent (BOLD) contrast related components generally entail low frequencies due to the timing of the hemodynamic response, components containing higher frequencies and components with time series correlating with realignment parameter time-series are classified as motion-related. The classification is also based on spatial features, and motion components are identified by high proportions of signal at the edge of the brain or in regions containing cerebrospinal fluid as head movement will make these regions appear as having strong signal variations even though this is due to voxels being located inside and then outside of the brain or vice versa when moved (Pruim et al., 2015). Components related to motion identified by AROMA were subsequently regressed out of the data.

5.2.2 Resting state connectivity analysis

HCP_MMP1.0 regions from Glasser et al. 2016 were mapped onto every participants' denoised resting state data. The fMRI signal time courses were extracted from our 11 regions of interest (LO1, LO2, LO3, V4t, MT, MST, FST, PHT, PH, TE1p, STSvp) and registered to MNI-space. Following Turken & Dronkers' (2011) analysis, a seed-to-whole-brain Pearson-correlation coefficient for time-courses in each voxel was calculated to explore the functional connectivity for each ROI in each participant. Maps of correlation values were Fisher-z-transformed and the resulting normally distributed z-values were submitted to one-sample t-tests in each voxel across participants separately for every ROI. This resulted in group-level maps containing t-values for each voxel indicating the functional connectivity of all eleven ROIs with the whole brain. To control for multiple comparisons, we used family-wise error rate across all brain voxels corresponding to $p < 0.001$ and a cluster-threshold of 40 contiguous voxels. We adopted the approach Turken & Dronkers (2011) used but applied a more intense noise cleaning using ICA-AROMA and a stricter threshold ($p < 0.001$ instead of $p < 0.01$) as we tested the functional connectivity of eleven different seed-ROIs.

The Pearson-correlation of the resulting connectivity map across the different regions was calculated to establish similarity of functional connectivity patterns of subregions. The clustering of the regions was tested using hierarchical clusters based on the unweighted average Euclidean distance and we depicted these with dendrograms. Figures were created using BrainNet Viewer (Xia et al., 2013, <http://www.nitrc.org/projects/bnv/>).

5.3 Results

The LOTC subregions were used as seed-ROIs for whole-brain resting state functional connectivity analyses and connectivity maps showed for both hemispheres networks including occipital, parietal, temporal and some also frontal regions. We illustrated functional connectivity maps in the ipsilateral hemisphere for seed ROIs in the left (Figure 41) and right hemisphere (Figure 42). Functional connectivity maps contralateral to the seed ROIs can be found in Appendix D.

For LOTC ROIs in both hemispheres we found strong correlations of time courses in the immediate vicinity of the ROIs for all subregions. Overall, seed-ROIs from both hemispheres revealed highly similar maps. For LO1 regions in both hemispheres, we found correlations of time courses in occipital cortex (inferior, lateral and superior occipital gyrus) extending into parieto-occipital sulcus. Some clusters were located around the central sulcus and the postcentral sulcus (SPL, IPL) and in the middle superior temporal sulcus. LO2 showed very similar connectivity maps but exhibited fewer and smaller clusters while for the LO3 seed, clusters in parieto-occipital sulcus, postcentral sulcus and central sulcus were stronger. Functional connectivity of V4t was similar to LO1, including a strong and extensive cluster in occipito-parietal cortex and smaller clusters in IPL, IPS and central sulcus. The same was found in MT and MST where we additionally found a cluster in posterior STG. FST again is connected to occipito-

parietal areas, but the cluster extends more into temporal cortex, SPL, IPL and central sulcus and exhibits a cluster in dIPMC, that was also found for MST, PH and PHT and a small cluster in vPMC and IFG in the left hemisphere. These clusters in dIPMC, vPMC and IFG increased in size and strength in the left hemisphere for PH, PHT and TE1p and also prefrontal regions exhibited similar time courses with these regions. In PH, we also found a strong cluster in posterior SPL, and while time courses in less occipital regions correlated, time courses from inferior temporal regions showed low correlation. In PHT, we find a map of broad extension with widespread clusters across frontal cortex, parietal cortex and a cluster in posterior temporal cortex. By contrast, we find confined clusters for TE1p in middle and inferior temporal cortex, angular gyrus and nearby regions in posterior parietal cortex and across posterior parts of frontal cortex extending though IFG into middle frontal gyrus. The most anterior seed-region STSvp revealed the smallest clusters of all regions located in superior temporal gyrus (one in the middle, one at the posterior end at the angular gyrus), middle frontal gyrus and prefrontal regions. Most of the regions in posterior and central LOTC were also functionally connected to clusters in lingual gyrus, cuneus and the cingulate sulcus, more central and anterior regions to precuneus, but for the sake of clarity, a medial view for the figure is included in Appendix D.

Overall, most connectivity maps of the posterior and central subregions of LOTC shared regions in occipital and parietal cortex, around the central sulcus, lingual gyrus and cuneus. Central regions of LOTC mainly connected to occipito-parietal cortex, central sulcus, postcentral sulcus, lingual gyrus, cuneus and cingulate sulcus. By contrast, subregions FST, PH and PHT also connected to frontal regions like dIPMC, vPC and IFG and extended more into temporal cortex. The anterior subregions revealed connected clusters close to the ones of posterior and central regions but the temporal cluster for TE1p was more anterior and for STSvp more anterior and superior, parietal clusters were more inferior at the angular gyrus and fewer frontal clusters were in inferior frontal gyrus but more in middle frontal gyrus.

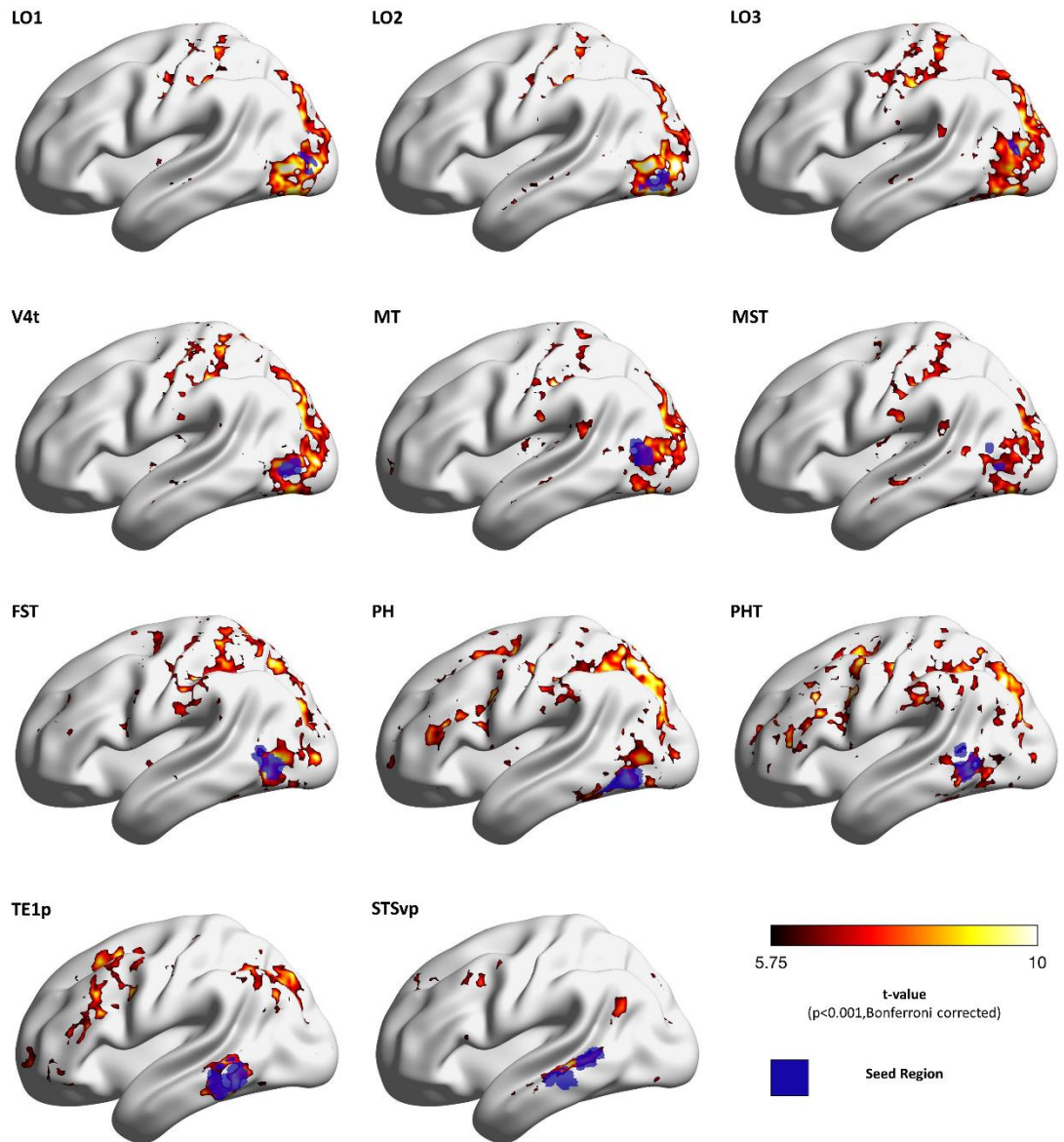


Figure 41. Left hemisphere whole brain functional connectivity maps from left hemisphere seed-ROIs in LOTC ($p < 0.001$, Bonferroni corrected, cluster size threshold of 40).

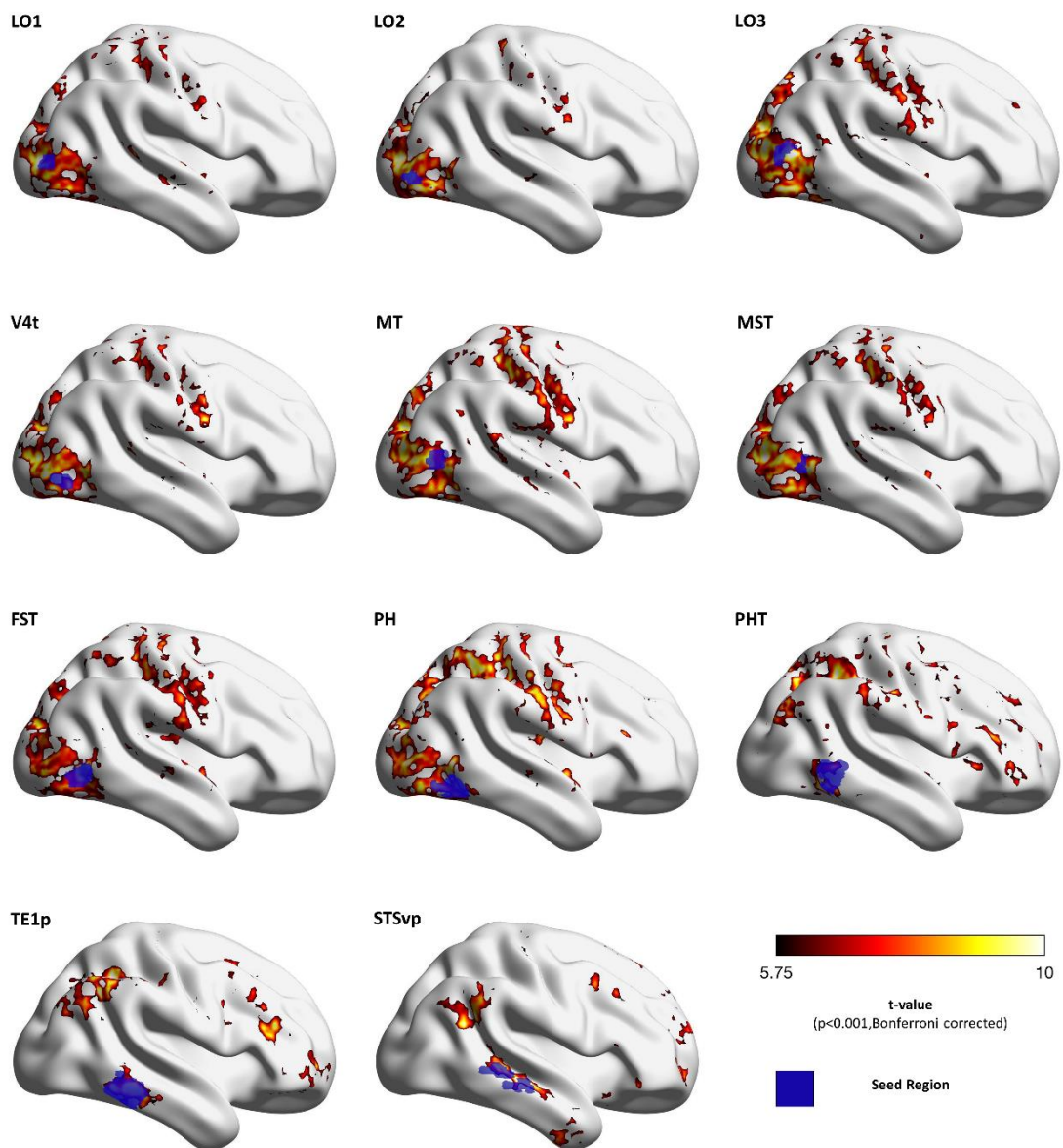


Figure 42. Right hemisphere whole brain functional connectivity maps from right hemisphere seed-ROIs in LOTC ($p < 0.001$, Bonferroni corrected, cluster size threshold of 40).

To evaluate the similarity of the different regions in more detail we used correlations of the connectivity maps from each seed-ROI and clustered the regions based on their resemblance (Figure 43). We find strong positive correlations between connectivity-maps of posterior and central subregions (LO1, LO2, LO3, V4t, MT, MST, FST) in LOTC, moderate correlations for central

regions PH and PHT, while the anterior regions TE1p and STSvp yield low correlations to most other subregions for seed-ROIs in both hemispheres (Figure 43, panel A). Connectivity maps from LO1, LO2, LO3, V4t, MT, MST, FST and in the right hemisphere also PH clustered together indicating similar connectivity maps for the posterior central regions of LOTC while a second and more anterior cluster was found consisting of TE1p, STSvp and PHT (in RH) unified by dissimilarity to other connectivity maps. In the left hemisphere, PH and PHT form a third cluster between the highly similar posterior/central regions and the highly dissimilar anterior regions by sharing moderate similarity to most other connectivity-maps (Figure 43, panel B).

The results also visualise hierarchies as regions close to each other exhibit stronger correlations than regions further apart (Figure 43, panel A). This is indicated by clusters close to the diagonal – as regions are sorted by their location and close by regions are also close by in the matrix. For example, MT is spatially closest to adjacent MST and LO3 but also their connectivity maps are clustered together closest.

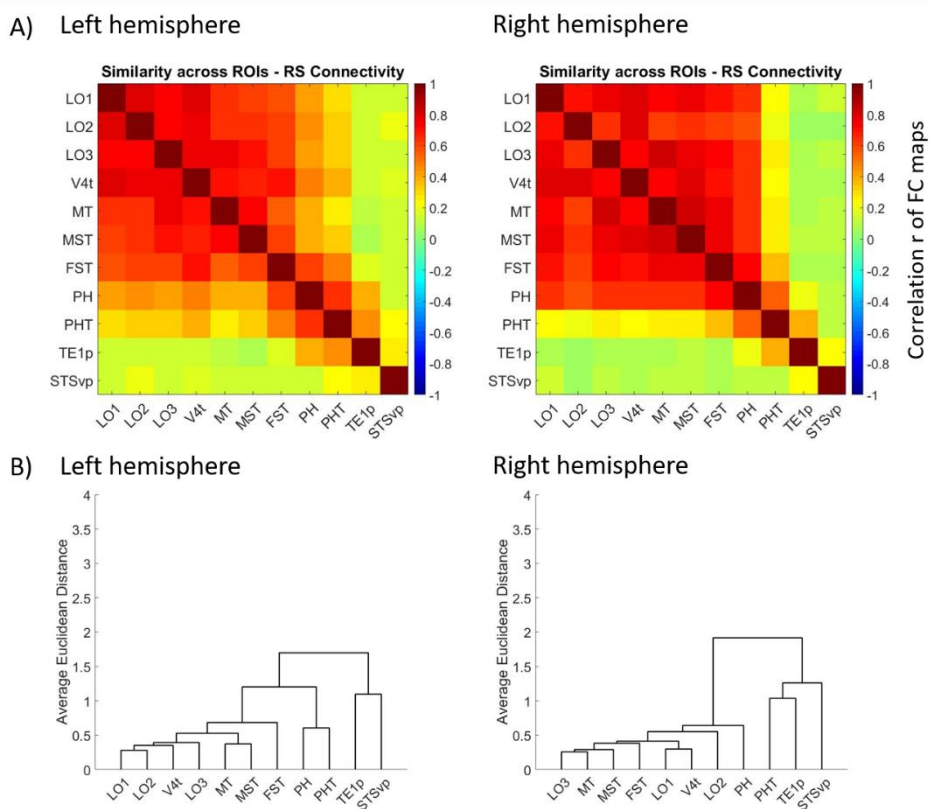


Figure 43. Similarity of functional connectivity (FC) maps across LOTC regions from left hemisphere seeds and right hemisphere seeds. A) Representational similarity matrix B) Dendrogram visualisation of results with hierarchical clustering.

5.4 Discussion

The analysis of the functional connectivity of subregions in LOTC revealed a main network of occipito-parietal regions, central and postcentral sulcus, lingual gyrus, cuneus, cingulate sulcus and superior temporal sulcus in the posterior and central subregions LO1, LO2, LO3, V4t, MT, MST and FST in both hemispheres. More anterior regions PH and PHT showed connectivity to similar regions but included also extensive clusters in frontal cortex and the occipito-parietal cluster was located more superior in the parietal regions and more anterior extending into temporal cortex. Surprisingly the networks related to the even more anterior subregions TE1p and STSvp were more localised and very different from the other networks including MTG,

angular gyrus, prefrontal cortex and middle frontal cortex for TE1p and STS, angular gyrus/posterior MTG, prefrontal regions and posterior parts of middle frontal gyrus.

The network related to posterior-central regions of LOTC is in line with what Turken and Dronkers (2011) reported for the MTG, where they also found similar connectivity to occipito-parietal and postcentral regions for the posterior section of MTG. We also find that more anterior regions of the LOTC PH, PHT and TE1p connect more to frontal regions and are more wide spread like they found for anterior sections of MTG. Surprisingly, we do not find this spread for the region STSvp, which exhibited only very few clusters and had less connection to frontal regions than the more posterior PHT.

Our results indicate two or possibly three divisions of LOTC corresponding to the previously reported segmentation from the representational similarity analysis with a posterior-central partition including LO1, LO2, LO3, V4t, MT, MST and FST, a central cluster of PH and PHT and the anterior division consisting of TE1p and STSvp. Studies investigating similar brain areas like the posterior lateral temporal cortex (Wang et al., 2015) or hubs of connectivity in the whole brain (Buckner et al., 2009) showed as well a separation of regions in the posterior temporal cortex and the middle and anterior regions in the temporal lobe.

In our results we find a hierarchical structure as well with regions that are located close to each other also having more similar networks, while the previous results showed stronger and more abrupt dissimilarities between the posterior-central and the anterior partitions.

A distributed coding of information in LOTC stands to reason as we did find only small and gradual changes in connectivity across the subregions we used, but we found different networks for posterior-central regions, central PH-PHT regions and for TE1p as well as for STSvp. These subdivisions could describe module-like regions as one feature of a module is that the contained neurons express similar properties – like their connection to other brain regions in this case.

However, the boundaries are not clear cut but rather gradual for the posterior-central and the PH-PHT divisions. A clear distinction can be seen for STSvp and TE1p though, indicating these regions are not part of a wider distributed subdivision that we suggest for the other LOTC regions

6 General Discussion

6.1 Chapter summary

This project aimed to advance the understanding of organizational principles underlying the neural responses of the LOTC to a range of localizer tasks. For this we evaluated functional MRI-data from tasks targeting regions sensitive to basic motion, biological motion, tools, body parts, bodies, action observation, action performance, verbs and spatial position as well as resting-state data from three different perspectives:

1. Spatial organization

We examined the spatial arrangement of responses with univariate analyses. Specifically, we evaluated average responses across participants, their overlap, the topography of peak activations and the selectivity of subregions in LOTC.

2. Representational organization

We investigated the representational array (multivariate analyses) by extracting which and how similarly tested domains are represented in LOTC subregions, how similar subregions are in terms of what they represent, and where and what information about spatial position is coded in or across LOTC.

3. Connectivity based organization

Finally, we employed functional connectivity of resting state data to see the array of connectivity of subregions in LOTC and how connectivity is similar across subregions.

The first central topic was the spatial organization of different domains within occipitotemporal cortex and possible subregions of LOTC. We replicated previous studies showing responses in LOTC for motion (Papeo & Lingnau, 2015; Tootell, 1995), biological motion (Papeo & Lingnau,

2015; Lingnau & Petris, 2012), action observation (Grezes & Decety, 2001; Caspers et al., 2010), action performance (Astafiev et al., 2004; Orlov et al., 2010), bodies (Downing et al., 2001), body parts (Orlov et al., 2010), tools (Chao et al., 1999) and verbs (Hafri et al., 2017; Bedny et al., 2008) and additionally found that these responses overlapped with each other substantially in the LOTC area. In contrast to previous research that examined these domains in separate studies and thus different groups of participants, by looking into the distribution of peak activation for all domains we found a reliable topography of the domains across our participants. Moreover, we detected overlapping confidence intervals for the domains located in more posterior portions of the LOTC, namely basic motion, biological motion, bodies, body parts and tools. By contrast, peak activations for action observation were obtained in more anterior and inferior portions of the LOTC, while peaks for action performance and verbs were found more anterior and superior. Both analyses revealed overlapping responses for posterior regions, speaking against an organization where the domains we tested are separated into modules, as modules would be expected to be separate functional entities selectively activating for only one domain. Testing this in more detail for subregions of LOTC, we found some regions, mainly located in posterior-central LOTC that showed high responses to several tasks including all domains but not the verb contrast. Anterior regions on the other hand showed weak responses overall and only for action observation, verbs and bodies the responses differed from zero. This shows again that posterior and central regions are probably not divided into domain-specific submodules – or at least not for the domains we used, as several subregions were not selective but rather undifferentiated, responding to several domains. A similar pattern of gradation of sensitivity instead of clear preference for one stimulus type was reported previously for fusiform face area, parahippocampal place area and right EBA (Downing, Chan, Peelen, Dots & Kanwisher, 2009), but comparing various object categories and scenes, Downing et al., (2009) also found EBA was not selective for bodies in the left hemisphere. Our analysis of the selectivity did not yield a clear

preference for one contrast, though the domains and stimuli we used included body related information on different levels, like whole bodies, body parts or bodies performing an action, and especially in the right hemisphere we also found strong responses to the stimuli visually presenting body related information. Bodies might be represented in LOTC on a more abstract level, invariant to the manner of presentation. The more anterior regions, however, only responded to few domains, indicating a better fit with a modular organization, but the overall low levels of activation and the small number of domains we obtained to be represented in anterior regions limit the certainty of this interpretation.

Repeatedly, we found a distinction of anterior regions and posterior-central regions in the occipito-temporal cortex where anterior parts show sparse responses to the more abstract domains of action observation, action performance and verbs while posterior-central areas respond strongly to a number of domains (basic motion, biological motion, tools, body parts, bodies but also action observation and performance as well). The results from the posterior-central division are in line with a distributed representation of several domains, sharing a representational space and hence inducing overlapping responses in the same area(s) instead of activating individual modules. As the anterior regions exhibit rather different results though, we suggest that our regions of interest cover different, confined representational spaces with different features represented. A tripartite organization of the ventral stream was suggested by Konkle and Caramazza (2008) looking into animacy and size as features. While they included regions more posterior to LOTC, the reported boundary between representations related to animal stimuli and small object stimuli in ventral stream might be located in the LOTC area. This further supports a division of regions in LOTC into subspaces and even though we did not use animals, objects or size as features, we still see a segmentation of posterior and anterior regions.

Representational organization is the topic of the third chapter where we extend the knowledge of the organization of LOTC by disentangling and clustering regions and domains based on the

pattern of activity they show in LOTC subregions. First, we checked which regions are sensitive to what domains and found most posterior-central regions to be sensitive to most domains except for verbs. This is in line with the meta-analysis from neurosynth in the first chapter and results from reviewed studies of regions recruited by the localizers (e.g. Papeo & Lingnau, 2015, Grossman et al., 2000, Bracci et al., 2011, Orlov et al., 2010, Downing et al., 2001; Caspers et al., 2010, Bedny et al., 2008). In the right hemisphere, none of the examined regions was recruited during the processing of verbs, in line with previously reported language lateralisation (e.g. Knecht et al., 2000), while in the left hemisphere sensitivity to verbs was mainly found in anterior regions. The sensitivity to all other domains drops abruptly at the PH and PHT subregions and no sensitivity was observable in anterior TE1p and STSvp for basic motion, biological motion, tools, body parts, bodies. Second, we evaluated how similarly the different domains are represented in the LOTC regions and found a body related cluster consisting of biological motion, body parts, bodies and action observation with similar representations in several subregions while the other domains basic motion, tools, action performance and verbs showed distinct patterns. Supporting this body cluster, a recent study reported LOTC to be sensitive to body-related features like the presence or orientation of a person (Wurm & Caramazza, 2019). EBA was previously related to features like body shape or posture (Downing & Peelen, 2011). Across the examined regions, these clustering results based on the similarity of the representations were quite consistent but again in anterior regions as the overall sensitivity to the domains is lost, the regions differ from posterior-central regions.

For the information of spatial position in the visual field we used a separate experiment and found very similar results regarding the posterior-central and anterior regions of LOTC as the representation of the hemifield location was found to be present in posterior-central regions but abruptly lost in anterior regions. These results reinforce the separation of posterior-central and anterior regions in LOTC and suggest that domains are not only overlapping and similarly

strong, but some tested domains are also represented similarly as indicated by the reported cluster of body-related contrasts. One important feature of the posterior-central representational space therefore might be the relation to the human body, but further research is needed to clarify what about the presence of a human body is represented in this cluster. The LOTC might be representing bodies not based solely on visual features like shape, parts or size (Bracci & Op de Beeck, 2016; Bracci et al., 2015; Konkle & Caramazza, 2013) but of a more holistic and abstract nature, generalizing across perceptual details like Op de Beeck, Pillet & Ritchie (2019) suggested this property of abstraction for occipito-temporal cortex. In LOTC, a hierarchical gradient of abstraction from exemplar specific information to representations that generalize across individual features like concrete shape or size to form a more categorical and abstract representation (Lingnau & Downing, 2015).

As basic motion, action performance and tools were represented in the regions as well but with different patterns of responses, the body feature might be sharing the region (or representational space) with additional, superimposed or related features like tool motion, (Beauchamp et al., 2002) or multimodality (integration of vision, action, and language information (Weiner & Grill-Spector, 2013), representing information important for action understanding.

The correlation of RSMs across subregions also made it more apparent, that regions located close to each other tend to have similar representational content and as subclusters form in close proximity to the diagonal indicating some gradual, possibly hierarchical principles underlying the organization of the information represented in LOTC.

Backing up the suggested two-part division of LOTC we found that posterior-central regions connected to a highly similar network while the anterior regions showed connections to separate and different networks. While the posterior-central regions mainly showed functional

connectivity with occipital and parietal regions, anterior regions tended to connect to frontal and temporal regions. Across the regions we found a diversification and extension of connections from posterior to central regions but a sudden reverse for anterior regions. As the subregions in posterior-central LOTC exhibited highly similar functional connections, this supports a shared utility with fine-grained, gradual differences while anterior regions form separate networks indicating different utility.

Comparing the similarity of categorical domain representation in the analysis of the action related localizers with the more gradual and specific analysis of similarity using the high level retinotopy data we find very similar RSMs and clusters with overall high similarity of representations within the clusters. The two-cluster solutions for the domain analysis as well as the HLR analysis are divided into posterior-central and anterior regions, but the domain RSM contains more differentiated subclusters while the HLR shows more a gradual, hierarchical pattern with increasing and decreasing similarity to and from the central cluster of MT, MST and FST. As the underlying data is very diverse for the domain experiments, for example looking at the similarity of motion, bodies or action observation a resulting richer subclustering makes sense. Numerous features like presence of bodies, motion or object size can be the reason for the subclustering. Op de Beeck et al. (2008) suggested a model of combined maps coding multiple functional properties in a set of voxels that is applicable to our findings as we used property-rich stimuli for this analysis. By contrast, we only investigated one property in the analysis in the HLR, which was the spatial location of the stimulus resulting in a gradual change of information as reported by Roth and Zohary (2015) and following a posterior to anterior gradient - this could describe just one of several properties in a combined map (Op de Beeck et al., 2008; Aflalo & Graziano, 2006); Graziano et al., 2005). Aflalo and Graziano (2006) generated an artificial motor cortex array to investigate several features relevant in motor cortex and the resulting map included somatotopic information, ethological motion categories and hand

location in 3D, represented in a hybrid map where all features influence the resulting topography. In LOTC, different features might also form a hybrid map influenced by different features to facilitate action understanding.

Comparing the results that investigate the similarity of subregions from the task based approach from chapter 4 with the results from the task-free resting state connectivity analysis in chapter 5, we found for both domain-localizers and high level retinotopy higher correlations between the individual subregions than in the task-free approach. Nevertheless, the clusters of regions were very similar for domain, HLR, and resting state results. The equivalence of overall functional architecture in resting state and task-based analysis (Smith et al., 2009) supports the assumption that there is a meaningful boundary between posterior-central and anterior regions of LOTC, separating cortical regions with highly different functional properties. Within these regions though, for task-based analysis the more fine-scaled sub-clustering was found with a gradual loss of similarity within clusters: highest consistency within subclusters was found for the domain-localizer experiments, less for the HLR experiment and lowest consistency within clusters for the resting-state data. This extends the previous argument of a multi-feature map (Op de Beeck et al., 2008; Aflalo & Graziano, 2006); Graziano et al., 2005) in LOTC representing different properties at the same time as we find gradually richer substructure within the regions with gradually richer stimulation.

The suggested parcellation of the LOTC in three subregions and their function in action perception are mainly based on functional fMRI using different tasks and stimuli. We have based the separation of regions and the possible function and underlying architecture on the basis of differences and similarities of representations of eight different tasks and twelve different contrasts. We designed the paradigms used to test different domains as similar as possible to each other as well as keeping them similar to the original work by using comparable block designs and a one-back task where applicable. Nevertheless, a limitation of the presented work

is that the comparison of unequal experiments and contrasts could show differences between the contrasts driven by differences in the stimulus sets, tasks or efficacy of the contrast. For example, the paradigm for basic motion has very simple stimuli with high similarity between the two conditions of moving and static dots, and the task was to only focus on the centre while the paradigm for action observation had rich and complex stimuli of action scenes, a dissimilar control condition of scrambled images and the task included categorization, remembering and decision about equality of the actions and a button press. The task is used to engage the participants in a specific cognitive process and processes of different complexity can only be evoked by different tasks and stimuli, like moving dots or pictures of action. To ensure that participants engage in, for example, action perception instead of pattern matching, a more complex task had to be used. Neuronal differences therefore could be affected by a number of differences between the experiments that we were not able to control for. This could have led to differences for example in the strength of the contrast driving comparably low t-scores for experiments with similar control (e.g. basic motion dots) conditions and high t-scores for less matched conditions (e.g. action observation and scrambled images). The size of the significant clusters and the selectivity might be affected by this, but for the clusters the main question was not how big but whether at all and where LOTC shows responses and only little clear-cut selectivity was found. For multivariate approaches, the level of t-values are not the key parameter but the underlying activation pattern. For example, the body-related tasks differed in terms of data sets, tasks and complexity and still clustered together. The overall results are not based on the absolute values but more the similarity of the location and strength of the contrasts despite the vast differences. Additionally, supporting the separation of the subregions in the LOTC, the results found for the task-based analyses were replicated with task-free resting state data where a systematic effect of different tasks and contrasts can be ruled out.

6.2 Future directions

The presented work gives rise to a great number of issues for future research that can be either tested with the existing comprehensive dataset or approached with future experiments. First of all, we did not use the full dataset yet and will be able to evaluate some domains in more detail. The body-related cluster in posterior-central LOTC for example could be evaluated in more detail, differentiating between different body-parts in the body-part localizer and between the different effectors used in the action performance localizer. As previous research reported different body-part topographies (Orlov, et al., 2010; Weiner & Grill-Spector, 2011), it would be interesting to evaluate whether this could also be a feature in the representational space and how similar the individual body parts are to whole bodies and the body related cluster we reported. How hands relate to tools more specifically can also be interesting as Bracci et al. (2012) found similar representations for tools and hands, but not tools, bodies or other body parts and as we did not find tools to be similar to body parts over all, a finer differentiation of body parts has to be considered.

Also not yet incorporated is the diffusion weighted imaging data and it would be interesting to compare functional connectivity with structural connectivity of LOTC subareas to see whether the white matter connections also align with a separation of posterior-central and anterior regions, and to what extent functional and structural boundaries correspond to each other.

Future studies could try to overcome the limitations of comparing diverse tasks, stimuli and contrasts with each other by designing a stimulus set where the different action components are not separated but combined in a naturalistic way like whole action scenes differing in key properties like implied motion, involved body parts or action category. As action perception is not a combination of separate properties but the features are related to each other – hands

usually are related to a body and feet and are often moving in a certain way—and the details of this interplay of features is important for the representation of the combination of the features as an action and should be investigated in the future. Moreover, different models could be built based on different features or combinations of features to evaluate what best explains neuronal responses and further the understanding of what the key properties for action perception are. Naturalistic, rich stimuli would also bring advantages like higher generalisability, experimental efficacy and validity and should be considered in the future.

The idea of a confined representational space in LOTC with different features represented by superimposed and possibly even additive maps also gives rise to numerous questions. First, where are the boundaries of this space? As our data indicated a sharp boundary between anterior and central LOTC, do we also find such a boundary at the superior, inferior or posterior end? Do all represented features show the same spatial restrictions, or are there differences between the size of the representational space they span – where some might be restricted to posterior regions while others extend to anterior regions as well? This would implicate that we know what features are represented though. Resolving this could be more difficult as additive superimposed features are difficult to disentangle and hardly any features are independent and orthogonalized (Felsen & Dan, 2005) but instead they often co-occur or correlate in the real world. For example, even though we used hand stimuli without a body, body stimuli without heads or actions without motion, in the real world these features coexist and are represented in our brain in concert. The real-world co-occurrence and correlation of proposed features represented in LOTC like spatial location (Kravitz, Vinson & Baker, 2008; Roth & Zohary, 2015), size, animacy (Konkle & Caramazza, 2013), or body properties (Orlov et al. 2010) might be represented accordingly in a cortical hub region like LOTC, as proposed by Lingnau & Downing (2015). The maps representing different features might not be several, independent maps, but instead align with the correlation of features with each other (Bracci, Ritchie & Op de Beeck,

2017). Further research is needed defining the dimensions of the feature space underlying this fascinating interaction and cooperation will be an important topic in the future.

6.3 Conclusion

In conclusion, we established a consistent boundary between anterior and posterior-central subregions of LOTC across investigated location, strength and pattern of responses as well as resting state functional connectivity. The posterior-central regions showed rich representations for several tested domains (basic motion, biological motion, tools, body parts, bodies, action observations and action performance) with shared representations for body-related domains. High similarity of strength and pattern of responses to most tasks across several regions suggests a distributed rather than a modular representation in a hub-region in posterior-central LOTC. Action understanding might be enabled by a hybrid map with superimposed or correlated feature maps in LOTC.

7 References

- Aflalo, T. N., & Graziano, M. S. (2006). Possible origins of the complex topographic organization of motor cortex: reduction of a multidimensional space onto a two-dimensional array. *Journal of Neuroscience*, *26*(23), 6288-6297.
- Amoruso, L., Couto, J. B., & Ibanez, A. (2011). Beyond Extrastriate Body Area (EBA) and Fusiform Body Area (FBA): context integration in the meaning of actions. *Frontiers in Human Neuroscience*, *5*, 124.
- Ariani, G., Wurm, M. F., & Lingnau, A. (2015). Decoding internally and externally driven movement plans. *Journal of Neuroscience*, *35*(42), 14160-14171.

- Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature neuroscience*, *7*(5), 542-548.
- Bajada, C. J., Trujillo-Barreto, N. J., Parker, G. J., Cloutman, L. L., & Ralph, M. A. L. (2019). A structural connectivity convergence zone in the ventral and anterior temporal lobes: Data-driven evidence from structural imaging. *cortex*, *120*, 298-307.
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, *34*(1), 149-159.
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of cognitive neuroscience*, *15*(7), 991-1001.
- Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., & Saxe, R. (2008). Concepts are more than percepts: the case of action verbs. *Journal of Neuroscience*, *28*(44), 11347-11353.
- Bedny, M., Dravida, S., & Saxe, R. (2014). Shindigs, brunches, and rodeos: The neural basis of event words. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(3), 891-901.
- Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological science*, *5*(4), 221-225.
- Birn, R. M., Murphy, K., & Bandettini, P. A. (2008). The effect of respiration variations on independent component analysis results of resting state functional connectivity. *Human brain mapping*, *29*(7), 740-750.
- Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. *Annu. Rev. Neurosci.*, *28*, 157-189.
- Bracci, S., Caramazza, A., & Peelen, M. V. (2015). Representational similarity of body parts in human occipitotemporal cortex. *Journal of Neuroscience*, *35*(38), 12977-12985.
- Bracci, S., Cavina-Pratesi, C., Connolly, J. D., & Ietswaart, M. (2016). Representational content of occipitotemporal and parietal tool areas. *Neuropsychologia*, *84*, 81-88.
- Bracci, S., Cavina-Pratesi, C., Ietswaart, M., Caramazza, A., & Peelen, M. V. (2012). Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *Journal of neurophysiology*, *107*(5), 1443-1456.
- Bracci, S., Ietswaart, M., Peelen, M. V., & Cavina-Pratesi, C. (2010). Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *Journal of Neurophysiology*, *103*(6), 3389-3397.
- Bracci, S., & Op de Beeck, H. (2016). Dissociations and associations between shape and category representations in the two visual pathways. *Journal of Neuroscience*, *36*(2), 432-444.

- Bracci, S., Ritchie, J. B., & Op de Beeck, H. (2017). On the partnership between neural representations of object categories and visual features in the ventral visual pathway. *Neuropsychologia*, *105*, 153-164.
- Brandman, T., & Yovel, G. (2014). Bodies are represented as wholes rather than their sum of parts in the occipital-temporal cortex. *Cerebral Cortex*, *26*(2), 530-543.
- Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nature reviews neuroscience*, *3*(3), 243.
- Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., ... & Johnson, K. A. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *Journal of neuroscience*, *29*(6), 1860-1873.
- Caramazza, A., & Shelton, J. R. (1998). Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of cognitive neuroscience*, *10*(1), 1-34.
- Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage*, *50*(3), 1148-1167.
- Chan, A. W., Kravitz, D. J., Truong, S., Arizpe, J., & Baker, C. I. (2010). Cortical representations of bodies and faces are strongest in commonly experienced configurations. *Nature neuroscience*, *13*(4), 417-418.
- Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature neuroscience*, *2*(10), 913-919.
- Culham, J., He, S., Dukelow, S., & Verstraten, F. A. (2001). Visual motion and the human brain: what has neuroimaging told us?. *Acta psychologica*, *107*(1-3), 69-94.
- de Haan, E. H., & Cowey, A. (2011). On the usefulness of 'what' and 'where' pathways in vision. *Trends in cognitive sciences*, *15*(10), 460-466.
- Deen, B., & McCarthy, G. (2010). Reading about the actions of others: biological motion imagery and action congruency influence brain activity. *Neuropsychologia*, *48*(6), 1607-1615.
- Downing, P. E., Chan, A. Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2005). Domain specificity in visual cortex. *Cerebral cortex*, *16*(10), 1453-1461.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470-2473.
- Downing, P., & Peelen, M. (2011). The role of occipitotemporal body-selective regions in person perception. *Cognitive Neuroscience*, *2* (3-4), 186-203.
- Downing, P. E., Peelen, M. V., Wiggett, A. J., & Tew, B. D. (2006). The role of the extrastriate body area in action perception. *Social Neuroscience*, *1*(1), 52-62.

- Downing, P. E., Wiggett, A. J., & Peelen, M. V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *Journal of Neuroscience*, *27*(1), 226-233.
- Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker Jr, C. L., Le Goualher, G., Pike, G. B., & Evans, A. C. (2000). A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cerebral cortex*, *10*(5), 454-463.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598.
- Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in cognitive sciences*, *18*(3), 120-126.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, *1*(1), 1-47.
- Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature neuroscience*, *8*(12), 1643.
- Ferri, S., Kolster, H., Jastorff, J., & Orban, G. A. (2013). The overlap of the EBA and the MT/V5 cluster. *Neuroimage*, *66*, 412-425.
- Ferrier, D. (1875). Experiments on the brain of monkeys. *Proceedings of the Royal Society of London*, *23*(156-163), 409-430.
- Fodor, J. (2001). Evolution and the Human Mind: Modularity, Language and Meta-Cognition Peter Carruthers and Andrew Chamberlin. *The British Journal for the Philosophy of Science*, *52*(3), 623-628.
- Fodor, J. A. (1983). *The modularity of mind*. MIT press.
- Fox, C. J., Iaria, G., & Barton, J. J. (2009). Defining the face processing network: optimization of the functional localizer in fMRI. *Human brain mapping*, *30*(5), 1637-1651.
- Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science*, *330*(6005), 845-851.
- Friston, K. J., Rotshtein, P., Geng, J. J., Sterzer, P., & Henson, R. N. (2006). A critique of functional localisers. *Neuroimage*, *30*(4), 1077-1087.
- Gallivan, J. P., McLean, D. A., Valyear, K. F., & Culham, J. C. (2013). Decoding the neural mechanisms of human tool use. *Elife*, *2*, e00425.
- Gallivan, J. P., McLean, D. A., Valyear, K. F., Pettypiece, C. E., & Culham, J. C. (2011). Decoding action intentions from preparatory brain activity in human parieto-frontal networks. *Journal of Neuroscience*, *31*(26), 9599-9610.

- Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2006). *Cognitive Neuroscience. The biology of the mind*, (2014). Norton: New York.
- Gerardin, E., Sirigu, A., Lehericy, S., Poline, J. B., Gaymard, B., Marsault, C., ... & Le Bihan, D. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cerebral cortex*, *10*(11), 1093-1104.
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., ... & Smith, S. M. (2016). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171-178.
- Golomb, J. D., & Kanwisher, N. (2011). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex*, *22*(12), 2794-2810.
- Grafton, S. T., & Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human movement science*, *26*(4), 590-616.
- Graziano, M. S., Aflalo, T. N., & Cooke, D. F. (2005). Arm movements evoked by electrical stimulation in the motor cortex of monkeys. *Journal of neurophysiology*, *94*(6), 4209-4223.
- Graziano, M. S., Taylor, C. S., & Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron*, *34*(5), 841-851.
- Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Human brain mapping*, *12*(1), 1-19.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annu. Rev. Neurosci.*, *27*, 649-677.
- Grosbras, M. H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human brain mapping*, *33*(2), 431-454.
- Grossman, E. D., Jardine, N. L., & Pyles, J. A. (2010). fMR-adaptation reveals invariant coding of biological motion on human STS. *Frontiers in human neuroscience*, *4*(15), 1-18.
- Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of cognitive neuroscience*, *12*(5), 711-720.
- Hafri, A., Trueswell, J. C., & Epstein, R. A. (2017). Neural representations of observed actions generalize across static and dynamic visual input. *Journal of Neuroscience*, *37*(11), 3056-3071.
- Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current opinion in Neurobiology*, *28*, 136-141.

- Hardwick, R. M., Caspers, S., Eickhoff, S. B., & Swinnen, S. P. (2018). Neural correlates of action: Comparing meta-analyses of imagery, observation, and execution. *Neuroscience & Biobehavioral Reviews*, *94*, 31-44.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*(3), 479-490.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*(2), 301-307.
- Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual review of neuroscience*, *37*, 435-456.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425-2430.
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., ... & Ramadge, P. J. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, *72*(2), 404-416.
- Hernández, M., Fairhall, S. L., Lenci, A., Baroni, M., & Caramazza, A. (2014). Predication drives verb cortical signatures. *Journal of cognitive neuroscience*, *26*(8), 1829-1839.
- Hock, H. S., & Nichols, D. F. (2013). The perception of object versus objectless motion. *Attention, Perception, & Psychophysics*, *75*(4), 726-737.
- Hodzic, A., Kaas, A., Muckli, L., Stirn, A., & Singer, W. (2009). Distinct cortical networks for the detection and identification of human body. *Neuroimage*, *45*(4), 1264-1271.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, *160*(1), 106-154.
- Hubel, D. H., & Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of neurophysiology*, *28*(2), 229-289.
- Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*(16), 7195-7205.
- Hutchison, R. M., Culham, J. C., Everling, S., Flanagan, J. R., & Gallivan, J. P. (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.

- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, *17*(2), 825-841.
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis*, *5*(2), 143-156.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & psychophysics*, *14*(2), 201-211.
- Kable, J. W., Lease-Spellmeyer, J., & Chatterjee, A. (2002). Neural substrates of action event knowledge. *Journal of Cognitive Neuroscience*, *14*(5), 795-805.
- Kanwisher, N. (2000). Domain specificity in face perception. *Nature neuroscience*, *3*(8), 759-763. Kanwisher, N. (2010). Functional specificity in the human brain: a window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences*, *107*(25), 11163-11170.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of neuroscience*, *17*(11), 4302-4311.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.
- Knecht, S., Deppe, M., Dräger, B., Bobe, L., Lohmann, H., Ringelstein, E. B., & Henningsen, H. (2000). Language lateralization in healthy right-handers. *Brain*, *123*(1), 74-81.
- Konen, C. S., & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature neuroscience*, *11*(2), 224-231.
- Konkle, T., & Caramazza, A. (2013). Tripartite organization of the ventral stream by animacy and object size. *Journal of Neuroscience*, *33*(25), 10235-10242.
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, *74*(6), 1114-1124.
- Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews Neuroscience*, *12*(4), 217.
- Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in cognitive sciences*, *17*(1), 26-49.
- Kravitz, D. J., Vinson, L. D., & Baker, C. I. (2008). How position dependent is visual object recognition?. *Trends in cognitive sciences*, *12*(3), 114-122.
- Kriegeskorte, N., & Bandettini, P. (2007). Combining the tools: activation-and information-based fMRI analysis. *Neuroimage*, *38*(4), 666-668.

- Kriegeskorte, N., Mur, M., & Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2, 4.
- Kühn, S., Keizer, A. W., Rombouts, S. A., & Hommel, B. (2011). The functional and neural mechanism of action preparation: roles of EBA and FFA in voluntary action control. *Journal of Cognitive Neuroscience*, 23(1), 214-220.
- Kwan, H. C., MacKay, W. A., Murphy, J. T., & Wong, Y. C. (1978). Spatial organization of precentral cortex in awake primates. II. Motor outputs. *Journal of neurophysiology*, 41(5), 1120-1131.
- Lambon Ralph, M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20120392.
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and emotion*, 24(8), 1377-1388.
- Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience*, 26(51), 13128-13142.
- Lewis, J. W. (2006). Cortical networks related to human use of tools. *The neuroscientist*, 12(3), 211-231.
- Lingnau, A., & Downing, P. E. (2015). The lateral occipitotemporal cortex in action. *Trends in cognitive sciences*, 19(5), 268-277.
- Lingnau, A., & Petris, S. (2012). Action understanding within and outside the motor system: the role of task difficulty. *Cerebral Cortex*, 23(6), 1342-1350.
- Lingnau, A., Ashida, H., Wall, M. B., & Smith, A. T. (2009). Speed encoding in human visual cortex revealed by fMRI adaptation. *Journal of Vision*, 9(13), 3-3.
- Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain?. *Trends in cognitive sciences*, 15(3), 97-103.
- Mahon, B. Z., Milleville, S. C., Negri, G. A., Rumiati, R. I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, 55(3), 507-520.
- Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in cognitive sciences*, 6(4), 176-184.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*, Henry Holt and co. Inc., New York, NY, 2(4.2).

- Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. (2011). Connectivity-based subdivisions of the human right “temporoparietal junction area”: evidence for different areas participating in different cortical networks. *Cerebral cortex*, *22*(8), 1894-1903.
- Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of neurophysiology*, *49*(5), 1127-1147.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of cognitive neuroscience*, *9*(5), 605-610.
- McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature reviews neuroscience*, *4*(4), 310.
- Mechelli, A., Crinion, J. T., Long, S., Friston, K. J., Ralph, M. A. L., Patterson, K., ... & Price, C. J. (2005). Dissociating reading processes on the basis of neuronal interactions. *Journal of cognitive neuroscience*, *17*(11), 1753-1765.
- Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI—an introductory guide. *Social cognitive and affective neuroscience*, *4*(1), 101-109.
- Nelissen, K., Borra, E., Gerbella, M., Rozzi, S., Luppino, G., Vanduffel, W., ... & Orban, G. A. (2011). Action observation circuits in the macaque monkey cortex. *Journal of Neuroscience*, *31*(10), 3743-3756.
- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, *395*(6705), 894.
- Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU Octave. *Frontiers in neuroinformatics*, *10*, 27.
- Op de Beeck, H., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience*, *26*(50), 13025-13036.
- Op de Beeck, H., Haushofer, J., & Kanwisher, N. G. (2008). Interpreting fMRI data: maps, modules and dimensions. *Nature Reviews Neuroscience*, *9*(2), 123.
- Op de Beeck, H., Pillot, I., & Ritchie, J. B. (2019). Factors Determining Where Category-Selective Areas Emerge in Visual Cortex. *Trends in cognitive sciences*.
- Orlov, T., Makin, T. R., & Zohary, E. (2010). Topographic representation of the human body in the occipitotemporal cortex. *Neuron*, *68*(3), 586-600.
- Orlov, T., Porat, Y., Makin, T. R., & Zohary, E. (2014). Hands in motion: an upper-limb-selective area in the occipitotemporal cortex shows sensitivity to viewed hand kinematics. *Journal of Neuroscience*, *34*(14), 4882-4895.

- Papeo, L., & Lingnau, A. (2015). First-person and third-person verbs in visual motion-perception regions. *Brain and language*, *141*, 135-141.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature reviews neuroscience*, *8*(8), 636.
- Peelen, M. V., Bracci, S., Lu, X., He, C., Caramazza, A., & Bi, Y. (2013). Tool selectivity in left occipitotemporal cortex develops without vision. *Journal of cognitive neuroscience*, *25*(8), 1225-1234.
- Peelen, M. V., Romagno, D., & Caramazza, A. (2012). Independent representations of verbs and actions in left lateral temporal cortex. *Journal of cognitive neuroscience*, *24*(10), 2096-2107.
- Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron*, *49*(6), 815-822.
- Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, *60*(4), 389-443.
- Petersen, S. E., & Sporns, O. (2015). Brain networks and cognitive architectures. *Neuron*, *88*(1), 207-219.
- Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *European Journal of Neuroscience*, *21*(10), 2864-2875.
- Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental brain research*, *209*(4), 481-493.
- Polimeni, J. R., Fischl, B., Greve, D. N., & Wald, L. L. (2010). Laminar analysis of 7 T BOLD using an imposed spatial activation pattern in human V1. *Neuroimage*, *52*(4), 1334-1346.
- Pruim, R. H., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *Neuroimage*, *112*, 267-277.
- Ramsey, R. (2018). Neural integration in body perception. *Journal of cognitive neuroscience*, *30*(10), 1442-1451.
- Reynaud, E., Lesourd, M., Navarro, J., & Osiurak, F. (2016). On the neurocognitive origins of human tool use: A critical review of neuroimaging data. *Neuroscience & Biobehavioral Reviews*, *64*, 421-437.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature neuroscience*, *2*(11), 1019.
- Ritchie, J. B., Bracci, S., & Op de Beeck, H. (2017). Avoiding illusory effects in representational similarity analysis: What (not) to do with the diagonal. *NeuroImage*, *148*, 197.

- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, 3(2), 131-141.
- Roth, Z. N., & Zohary, E. (2015). Position and identity information available in fMRI patterns of activity in human visual cortex. *Journal of Neuroscience*, 35(33), 11559-11571.
- Samartsidis, P., Montagna, S., Nichols, T. E., & Johnson, T. D. (2017). The coordinate-based meta-analysis of neuroimaging data. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 32(4), 580.
- Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain*, 130(9), 2452-2461.
- Saygin, A. P., McCullough, S., Alac, M., & Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral temporal cortex by real and fictive motion sentences. *Journal of cognitive neuroscience*, 22(11), 2480-2490.
- Saygin, A. P., Wilson, S. M., Hagler, D. J., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *Journal of Neuroscience*, 24(27), 6181-6188.
- Schwarzbach, J. (2011). A simple framework (ASF) for behavioral and neuroimaging experiments based on the psychophysics toolbox for MATLAB. *Behavior research methods*, 43(4), 1194-1201.
- Sereno, M. I., Dale, A. M., Reppas, J. B., & Kwong, K. K. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268(5212), 889.
- Sereno, M. I., McDonald, C. T., & Allman, J. M. (1994). Analysis of retinotopic maps in extrastriate cortex. *Cerebral Cortex*, 4(6), 601-620.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human brain mapping*, 17(3), 143-155.
- Smith, S. M. (2004). Overview of fMRI analysis. *The British Journal of Radiology*, 77(suppl_2), S167-S175.
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., ... & Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, 106(31), 13040-13045.
- Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, 44(1), 83-98.
- Sporns, O., & Betzel, R. F. (2016). Modular brain networks. *Annual review of psychology*, 67, 613-640.

- Starck, T., Nikkinen, J., Rahko, J., Remes, J., Hurtig, T., Haapsamo, H., ... & Pauls, D. L. (2013). Resting state fMRI reveals a default mode dissociation between retrosplenial and medial prefrontal subnetworks in ASD despite motion scrubbing. *Frontiers in human neuroscience*, *7*, 802.
- Tarhan, L., & Konkle, T. (2019). Sociality and Interaction Envelope Organize Visual Action Representations. *bioRxiv*, 618272.
- Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science*, *352*(6282), 216-220.
- Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., ... & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, *15*(4), 3215-3230.
- Tucciarelli, R., Wurm, M., Baccolo, E., & Lingnau, A. (2019). The representational space of observed actions. *eLife*, *8*, e47686.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage*, *16*(3), 765-780.
- Turken, U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in systems neuroscience*, *5*.
- van Boxtel, J. J., & Lu, H. (2013). A biological motion toolbox for reading, displaying, and manipulating motion capture data in research settings. *Journal of vision*, *13*(12), 7-7.
- Van Den Heuvel, M. P., & Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *European neuropsychopharmacology*, *20*(8), 519-534.
- Vannuscorps, G., Wurm, M. F., Striem-Amit, E., & Caramazza, A. (2018). Large-Scale Organization of the Hand Action Observation Network in Individuals Born Without Hands. *Cerebral Cortex*, *1*, 11.
- Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 625-629.
- Vernon, R. J., Gouws, A. D., Lawrence, S. J., Wade, A. R., & Morland, A. B. (2016). Multivariate patterns in the human object-processing pathway reveal a shift from retinotopic to shape curvature representations in lateral occipital areas, LO-1 and LO-2. *Journal of Neuroscience*, *36*(21), 5763-5774.
- Wager, T. D., Lindquist, M., & Kaplan, L. (2007). Meta-analysis of functional neuroimaging data: current and future directions. *Social cognitive and affective neuroscience*, *2*(2), 150-158.

- Wang, X., Peelen, M. V., Han, Z., He, C., Caramazza, A., & Bi, Y. (2015). How visual is the visual cortex? Comparing connectional and functional fingerprints between congenitally blind and sighted individuals. *Journal of Neuroscience*, *35*(36), 12545-12559.
- Watson, C. E., & Buxbaum, L. J. (2015). A distributed network critical for selecting among tool-directed actions. *Cortex*, *65*, 65-82.
- Watson, C. E., Cardillo, E. R., Ianni, G. R., & Chatterjee, A. (2013). Action concepts in the brain: an activation likelihood estimation meta-analysis. *Journal of cognitive neuroscience*, *25*(8), 1191-1205.
- Weiner, K. S., & Grill-Spector, K. (2011). Not one extrastriate body area: using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *Neuroimage*, *56*(4), 2183-2199.
- Weiner, K. S., & Grill-Spector, K. (2013). Neural representations of faces and limbs neighbor in human high-level visual cortex: evidence for a new organization principle. *Psychological research*, *77*(1), 74-97.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). *Controlling low-level image properties: the SHINE toolbox. Behavior research methods*, *42*(3), 671-684.
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *Neuroimage*, *92*, 381-397.
- Wu, J., Ngo, G. H., Greve, D., Li, J., He, T., Fischl, B., ... & Yeo, B. T. (2018). Accurate nonlinear mapping between MNI volumetric and FreeSurfer surface coordinate systems. *Human brain mapping*, *39*(9), 3793-3808.
- Wurm, M. F., & Caramazza, A. (2019). Distinct roles of temporal and frontoparietal cortex in representing actions across vision and language. *Nature communications*, *10*(1), 289.
- Wurm, M. F., & Lingnau, A. (2015). Decoding actions at different levels of abstraction. *Journal of Neuroscience*, *35*(20), 7727-7735.
- Wurm, M. F., Caramazza, A., & Lingnau, A. (2017). Action categories in lateral occipitotemporal cortex are organized along sociality and transitivity. *Journal of Neuroscience*, *37*(3), 562-575.
- Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS one*, *8*(7), e68910.
- Yarkoni, T. (2011). Neurosynth.org. from neurosynth.org.
- Yarkoni, T. (2017). Neurosynth core tools v0. 3.1.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991).
A direct demonstration of functional specialization in human visual cortex.
Journal of neuroscience, 11(3), 641-649.

8 Appendices

Appendix A: Instructions for experiments

Instructions for the LOTC Study

Dear Participant

Welcome to the Lab and thank you for taking part in this study!

This study includes eight different experiments with simple tasks, taking place on two separate days. Moreover, we will run a number of sequences during which no specific tasks are required. During these sequences, we will simply ask you to stay still and relax. At the end of both scanning sessions, we will provide you with a brief questionnaire. Each of the two scanning sessions will last approximately 75 min. Here is a brief overview of the two scanning sessions:

Day 1

Scout (no task)	30 sec
Task 1 (Videos)	11 min
Task 2 (Words)	17 min
Anatomical scan (no task)	6 min
opp. phase encode (no task)	30 sec
Field mapping (no task)	2 min
Task 3 (Motion)	11 min
Task 4 (Movement Execution)	9 min
Resting state (no task; eyes closed)	10 min

Day 2

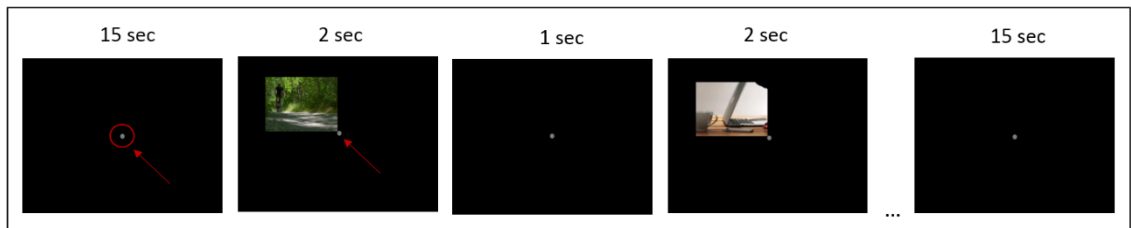
Scout (no task)	30 sec
Task 5 (Animals Tools Bodies Chairs)	21 min
Task 6 (Body Parts)	11 min
Task 7 (Biological Motion)	11 min
Task 8 (Action Observation)	9 min
DTI (no task)	9 min

All tasks are explained in these written instructions. Before entering the MRI lab, you will be able to familiarize yourself with the tasks before the scanning session.

Day 1

Task 1: Videos

In this experiment, different videos of actions will be shown either in the upper left, upper right, lower left, or lower right quadrant of the screen. Please fixate on the dot in the centre throughout the experiment. Once in a while, the fixation dot will briefly become a bit darker. Your task will be to press a button with your right index finger as soon as you detect this change.

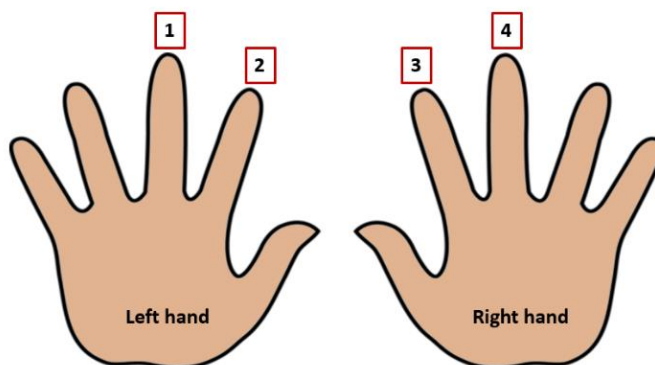
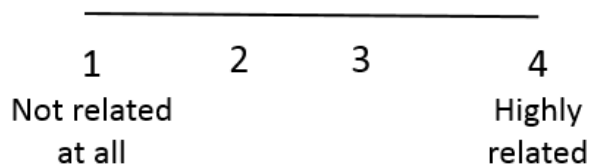
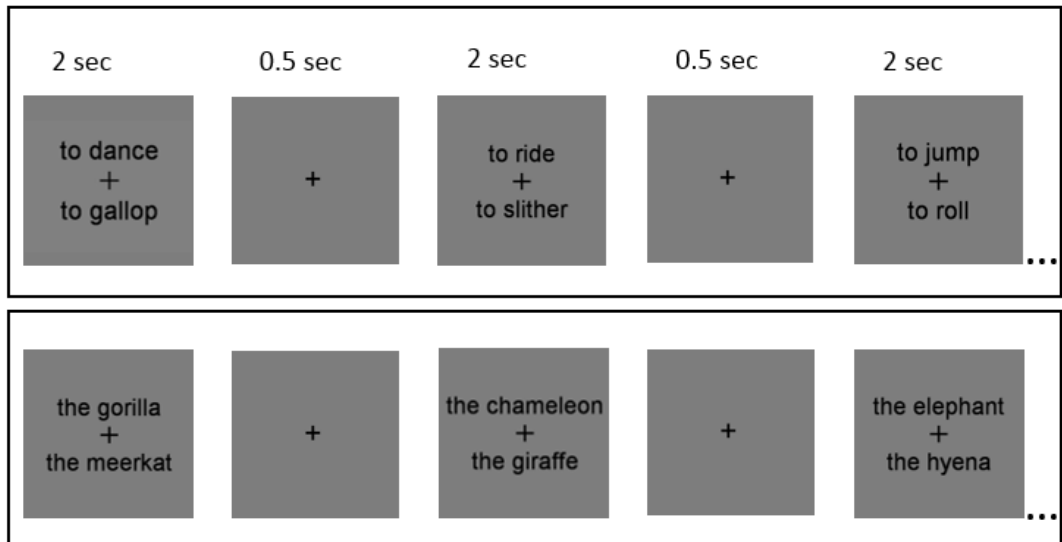


One run of this experiment will take about 5:15 minutes and there will be two runs (10:30 min) in total.

Day 1

Task 2: Words

In this experiment you will be shown two words at a time. If words are present, you do not have to fixate in the centre, but you can read the words. Your task is to indicate how related in meaning the words are on a scale from one to four (1- not related at all- 4 highly related) by pressing one of four buttons with your left and right middle and index fingers. Please try to respond while the words are on the screen, or during the short break (0.5 sec) afterwards.

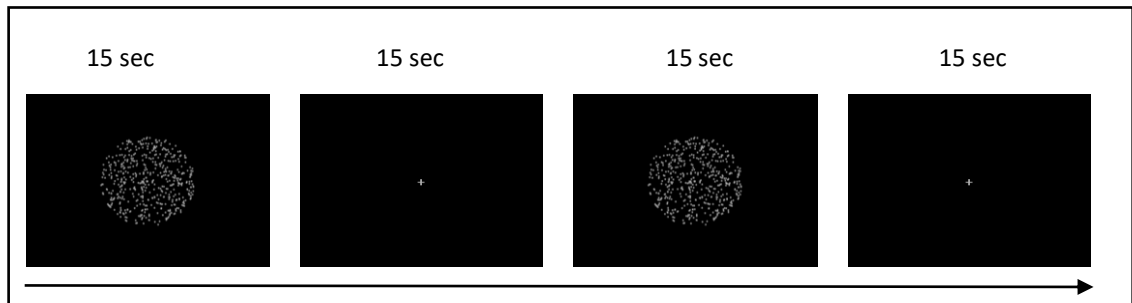


One run of this experiment will take about 4:15 minutes and there will be four runs (17 min) in total.

Day 1

Task 3: Motion

In this experiment, you will be shown white dots on a black background. The dots will be sometimes moving and sometimes be static. During this task, please focus on the dot and the fixation cross in the centre of the screen. You do not need to respond during this task.



One run of this experiment will take 5.15 minutes and there will be two runs (10.30 min) in total. After each run you can have a short break.

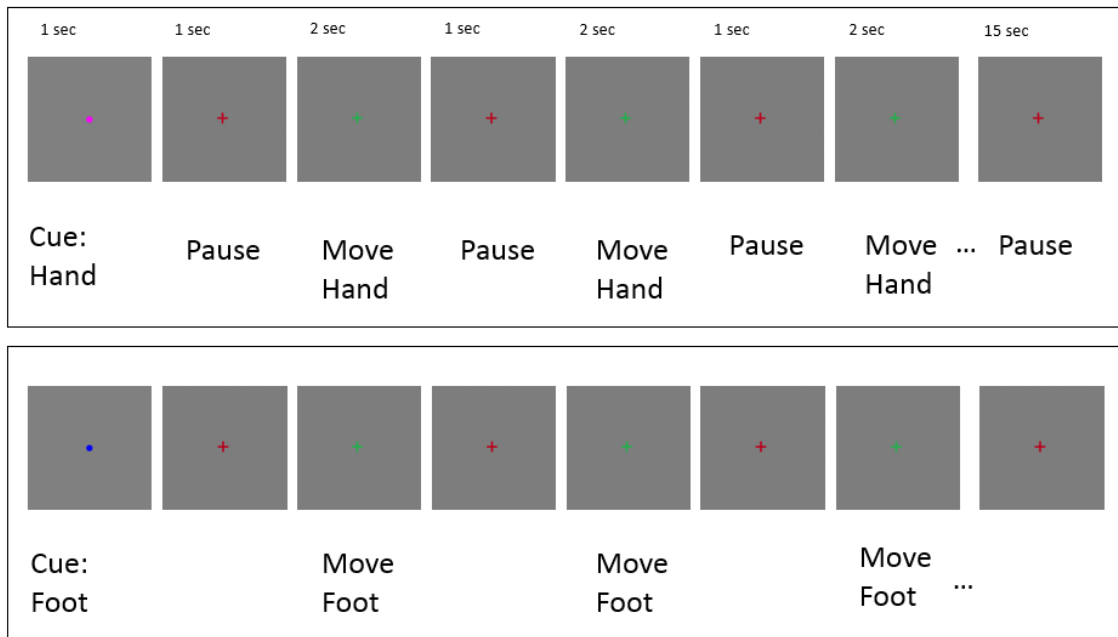
Day 1

Task 4: Movement execution

Even though we would like to ask you to minimize motion, especially of your head throughout the scan, in this experiment you will be asked to move either your right hand or right foot.

This experiment consists of 15 second blocks during which you are supposed to move your right hand or foot, interleaved with 15 second blocks during which you can rest. At the beginning of each movement block, you will be presented with a coloured disk that indicates which body part to move. If the disk is pink, in the upcoming block please move your hand as if you want to turn on the light by pressing a light switch. If the disk at the beginning of a block is blue, please move your foot as if you would turn on a standard lamp with a foot operated switch on the floor when the cross turns green.

Within each movement block, the fixation cross will turn green whenever you are supposed to make a movement, while a red fixation cross indicates a break.



One run of this experiment will take 4.15 minutes and there will be two runs (8.3 min) in total.

Day 2

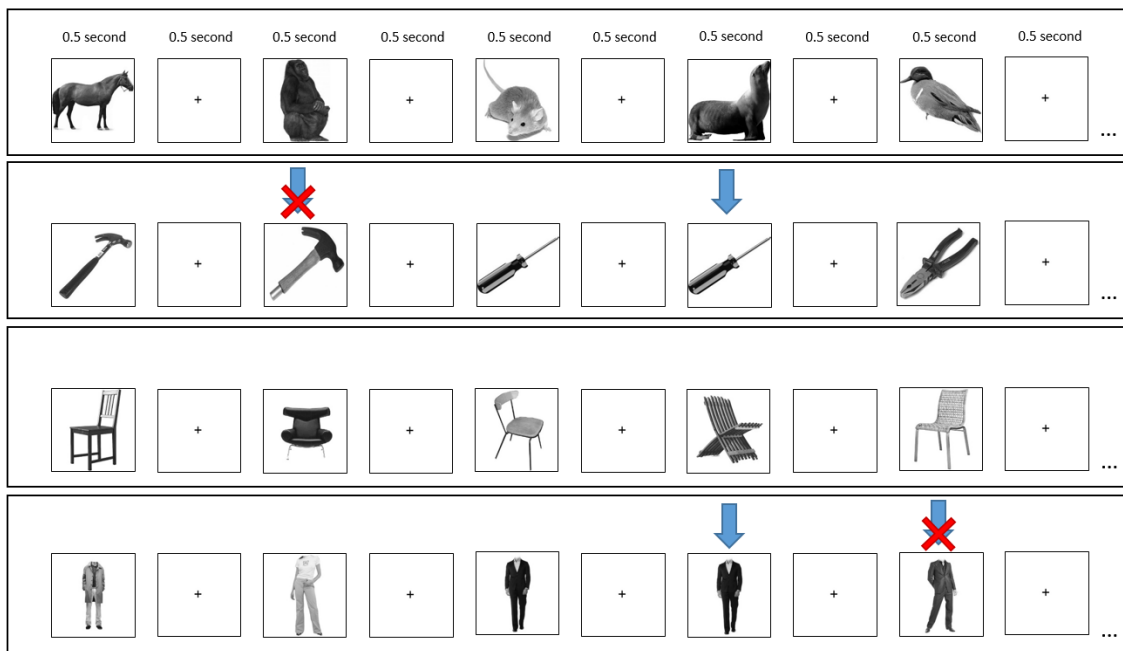
Task 5: Animals, Tools, Bodies, and Chairs

In this experiment, pictures of bodies, tools, animals & chairs will be presented. Your task is to indicate with a button press with your right index finger whenever you see a consecutive repetition of the identical image (see example). Some pictures can be very similar but are not the same – for example two different hammers (see example row 2). Only press the button if it is the exact same exemplar shown successively, like the screwdriver. Please give your response while the image is on the screen or during the short break (0.5 sec) afterwards.

If there is no repetition for a while, do not worry, it happens infrequently.

Example picture task 1:

Repetition indicated by the blue arrow – press button during or right after this picture – not during or after pictures where the arrow is crossed out like in the second row where the same object is shown, a hammer, but different exemplars.



One run of this experiment will take about 5:15 minutes and there will be four runs (21 min) in total.

Day 2

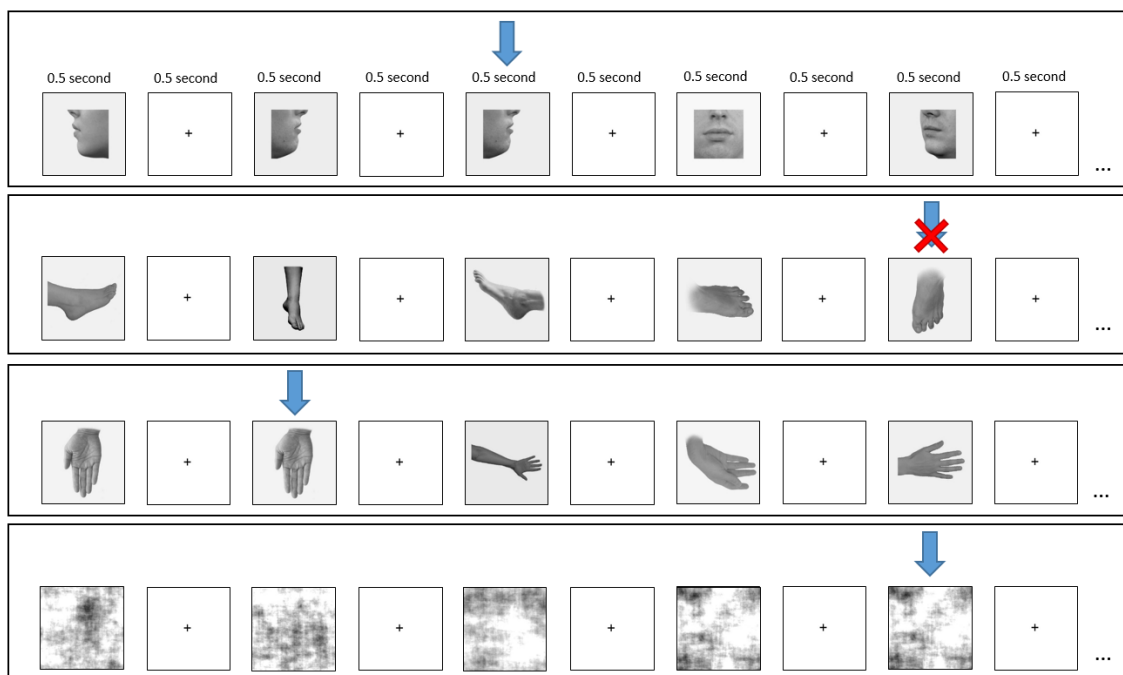
Task 6: Body Parts

In this experiment pictures of body parts and scrambled pictures will be shown. Your task is to indicate with a button press of the right index finger whenever you see a direct repetition of the same exemplar. Some pictures can be very similar but are not the same – for example the same foot from a different angle (see example row 2). Only press the button if it is the exact same exemplar shown successively. You can respond during the body part and the short break (0.5 sec) afterwards. The next picture will automatically be shown after half a second regardless whether there is a button press or not.

If there is no repetition for a while, do not worry, it happens infrequently.

Example picture task 1:

Repetition indicated by the blue arrow – press button during or right after this picture – not during or after pictures where the arrow is crossed out like in the second row where the same foot is shown, but an altered exemplar is presented where the orientation differs.



One run of this experiment will take 5:15 minutes and there will be two runs (10:30 min) in total.

Day 2

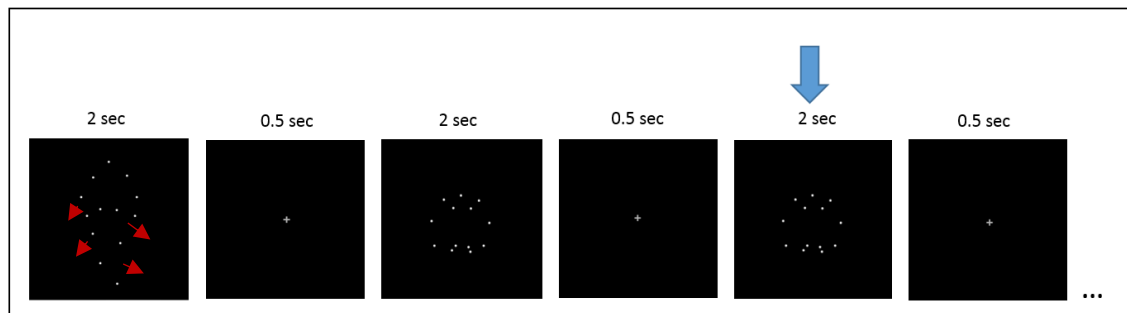
Task 7: Point Light Actions

This experiment consists of white dots on black background depicting actions like walking or cycling. Sometimes you will see the white dots moving without depicting a specific action. Short videos of actions or moving dots will be shown, followed by a short break indicated by a fixation cross. Occasionally the same video will be shown twice consecutively. Your task is to indicate with a button press whenever you see a repetition of the same exemplar.

Please try to respond during the video and the short break (0.5 sec) afterwards. You can see the actual actions later on during the familiarization at the PC, but here is already an outline of the task. If there is no repetition for a while, do not worry, it happens infrequently.

Example point light actions:

Repetition indicated by the blue arrow – press button during or right after this video.



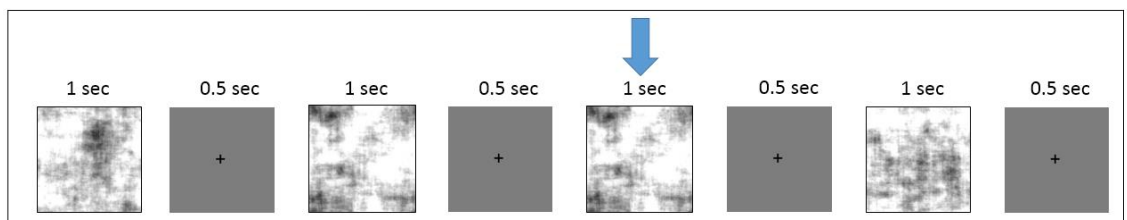
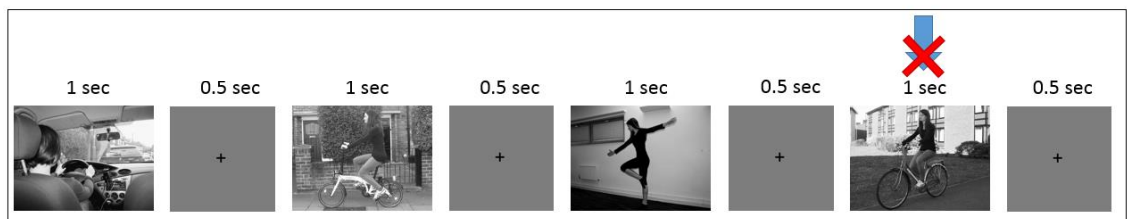
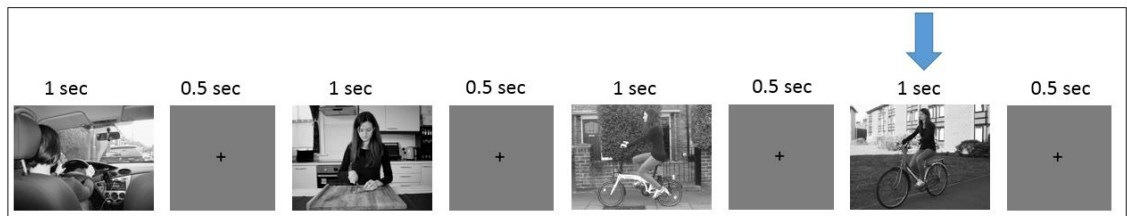
One run of this experiment will take 4.15 minutes and there will be two runs (8.3 min) in total.

Day 2

Task 8: Action observation

In this experiment, you will be presented with pictures depicting different kinds of actions such as running, eating or hugging. Once in a while, the same action (e.g. riding the bicycle) will be shown in two consecutive trials (see example in the top row of the following picture). Your task will be to press a button whenever you see a repetition of an action. In between there will be pictures of grey and white patterns. Your task is to indicate with a button press of the right index finger whenever you see a direct repetition of the same exemplar of these patterns. Please note that you should press a button only when the repetition takes place in two consecutive trials (see example in the bottom row of the following picture: same action –riding a bike- but not in two consecutive trials). Please try to respond during the picture and the short break (0.5 sec) afterwards.

If there is no repetition for a while, do not worry, it happens infrequently.



One run of this experiment will take about 4:15 minutes and there will be two runs (8.3 min) in total.

Appendix B:Post session questionnaire

LOTC • Post-session questionnaire

BEFORE the experiment started

- Did you have coffee right before coming to the MR lab? (YES/ NO)

- Were the instructions and training sufficient before entering the scanner? (YES/ NO)

DURING the experiment

- How comfortable did you feel once inside the scanner?

1 2 3 4 5 6 (from 1 = *very uncomfortable*, to 6 = *very comfortable*)

- How tired did you get throughout the experimental session?

1 2 3 4 5 6 (from 1 = *not at all*, to 6 = *falling asleep*)

- How much did you internally verbalize the actions represented in the pictures?

1 2 3 4 5 6 (from 1 = *not at all*, to 6 = *quietly naming/lips moving*)

- How much were you concentrated exclusively on the repetition of the action category?

1 2 3 4 5 6 (from 1 = *not concentrated at all*, to 6 = *concentrated exclusively on that*)

- Did you make use of any specific strategy to handle the task? If yes, could you please describe which?

AFTER the experiment

- Did you ever feel any kind of distress? Do you have any suggestions for improvements for future sessions?

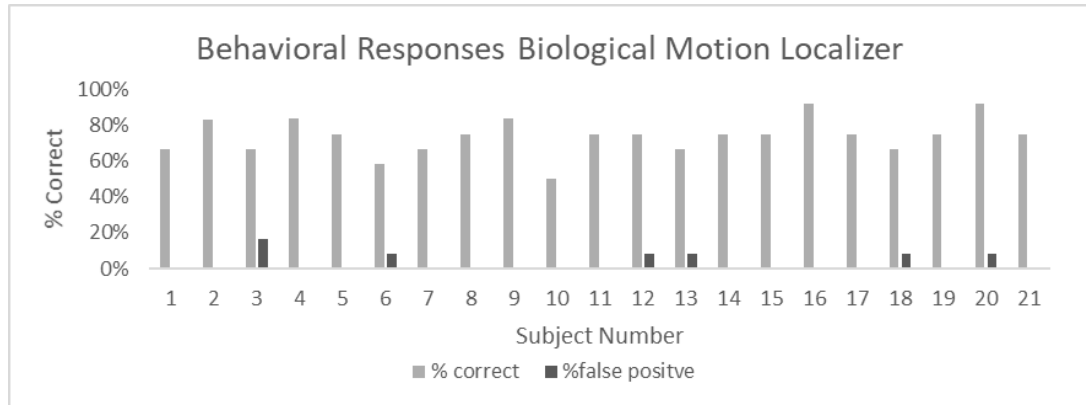
Thanks!

- Do you have any additional comments/general remarks about the experiment?

Appendix C: Behavioral Results

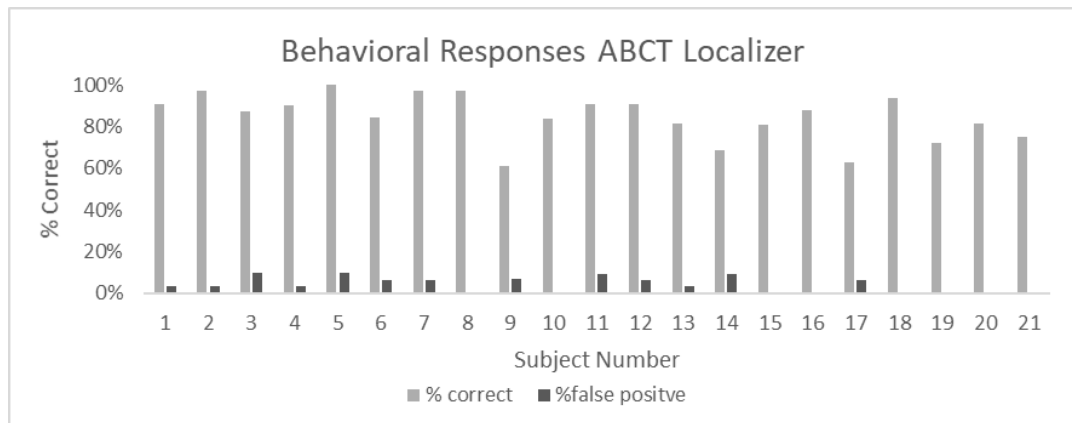
Biological Motion

Responses for all participants showed that all participants performed the task with a mean accuracy of 74% (SD = 10%) and a mean rate of false positive responses of 3%.



Bodies & Tools

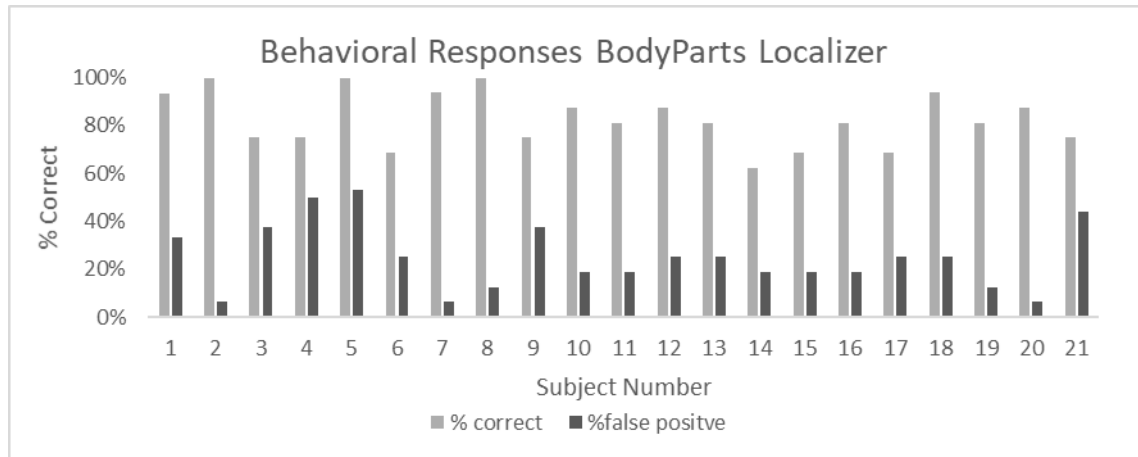
For the ABCT Localizer, participants responded to 84% (SD=11%) of the trials correctly and indicated only in 4% a repetition, when there was none.



Body Parts

The average for correct responses in this localizer is at 72% (SD= 14%) with a mean of 30% false positive responses. The high false positive rate might be explained by a high similarity of stimuli

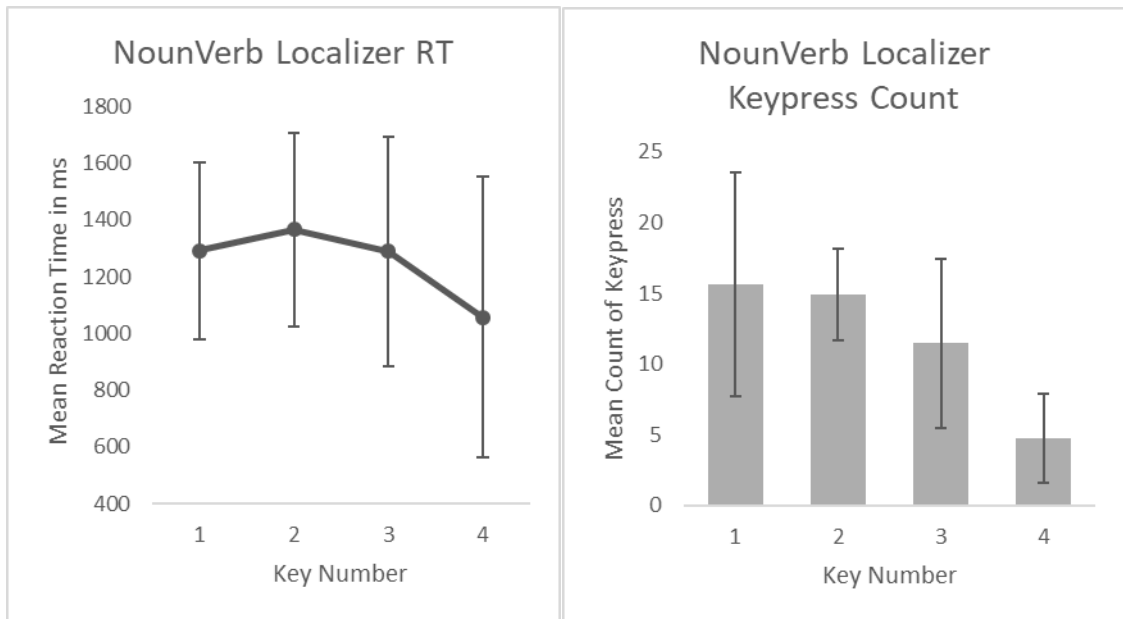
of the same category sometimes – the same hand in a different orientation or two different hands in a comparable posture look very similar but would not be a repetition of the exact same stimulus.



Verbs

An evaluation of the response time and frequency of each of the response options (1-4 with 1= words are not related at all to 4= words are highly related in meaning) showed a mean reaction time of 1250.7 ms. A repeated measures ANOVA with a post-hoc test indicated an overall effect of response option on the reaction time ($F(1.99,39.78) = 7.95$; $p=0.001$). This effect is driven by the faster responses for the 4th key (very related in meaning) compared to keys 2 ($p<0.001$) and 3 ($p=0.005$). No other significant differences between all other response options were found.

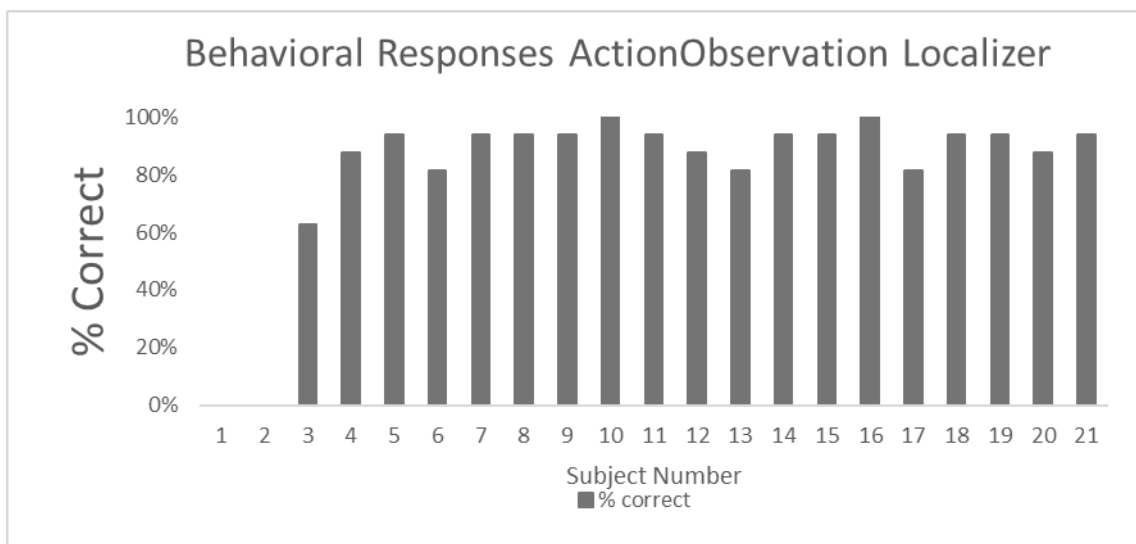
The count of each response option indicated that the option 1 (not related at all) was chosen 15.64 times on average, key 2 14.91 times, key 3 11.45 times and key 4 was chosen 4.74 times on average. A comparison with a repeated measures ANOVA highlighted, that the 4th key is chosen significantly less often than all other options (1&4: $p<0.001$;2&4: $p<0.001$; 3&4: $p<0.001$), driving the overall effect. Options 1, 2 and 3 were chosen with a similar frequency.



Mean reaction time for response keys 1 (not related) -4 (highly related) and mean count of response keys. Error bars represent SD.

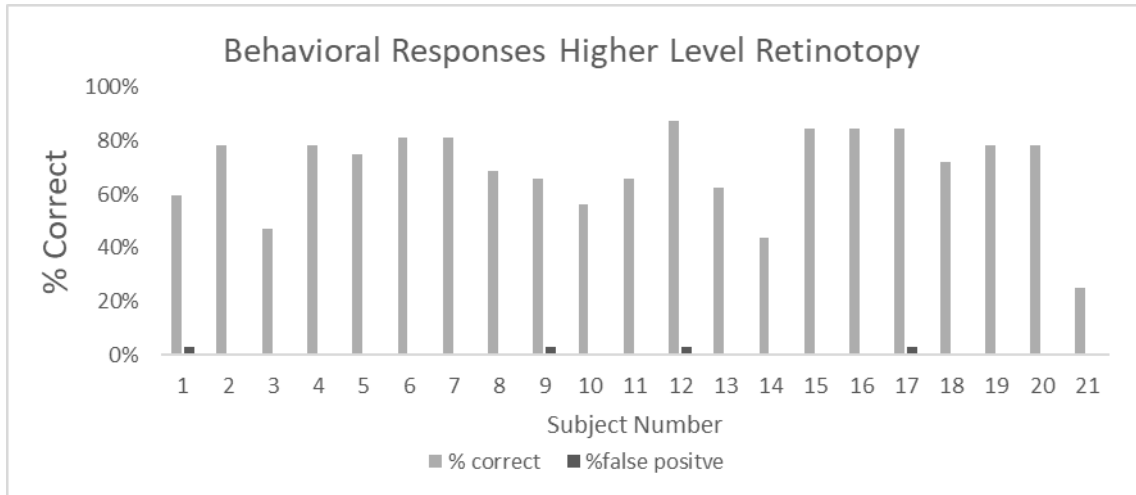
Action Observation

Participants detected most of the repetitions of the action observation localizer (M=90% correct, SD=9%).



Higher Level Retinotopy

Overall 69% (SD= 16%) correct responses were given during this task, only 1% false positive responses.



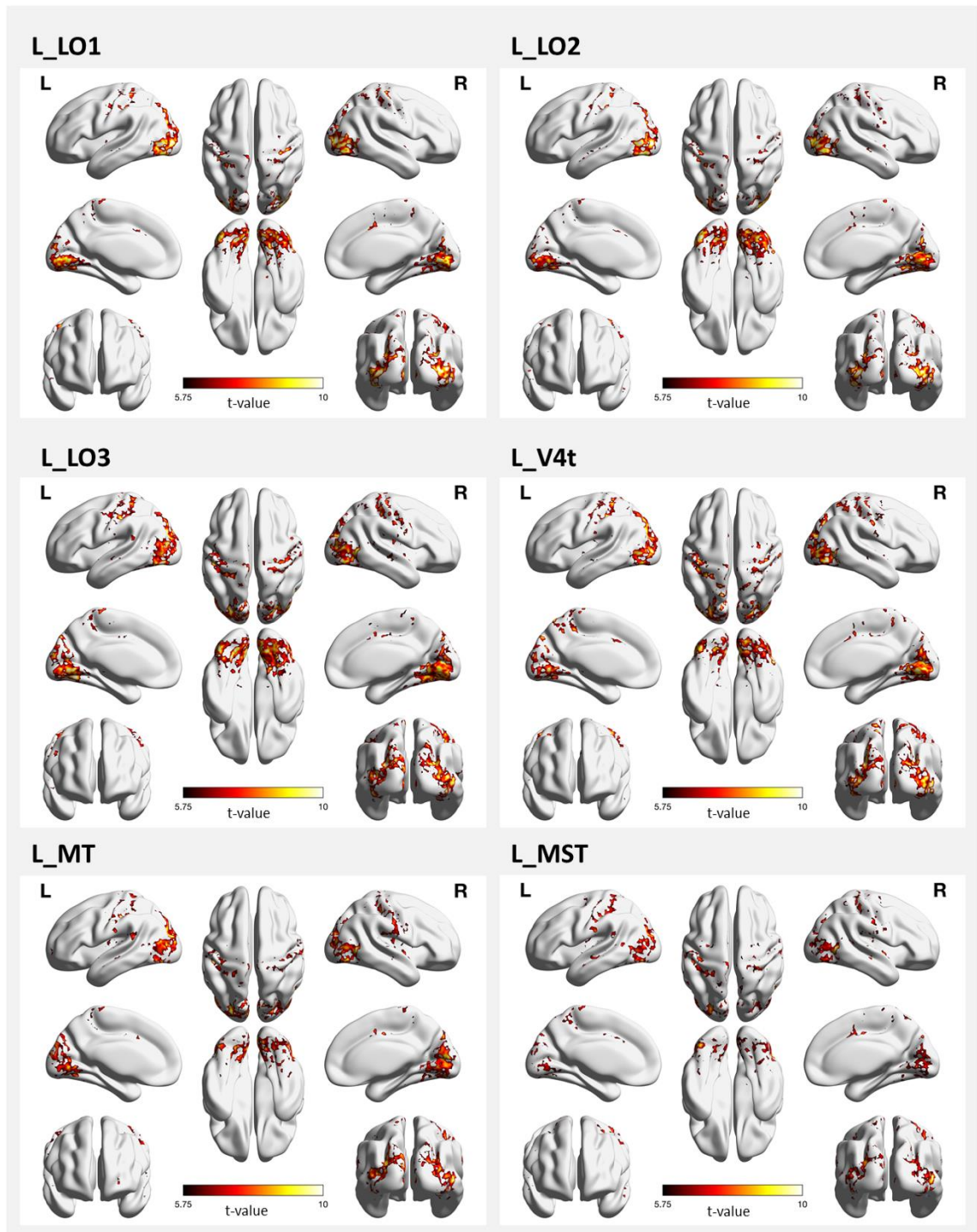
Basic Motion

Participants were only asked to fixate at the central dot on the screen and no responses were collected.

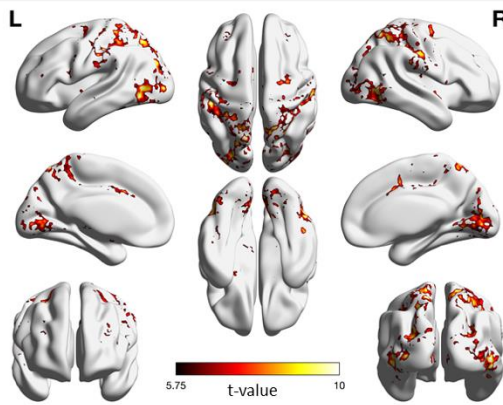
Movement Execution

No responses were collected for this localizer.

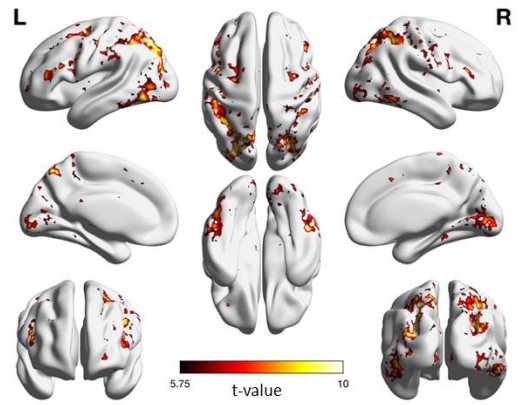
Appendix D: Full view results for functional seed-whole brain connectivity



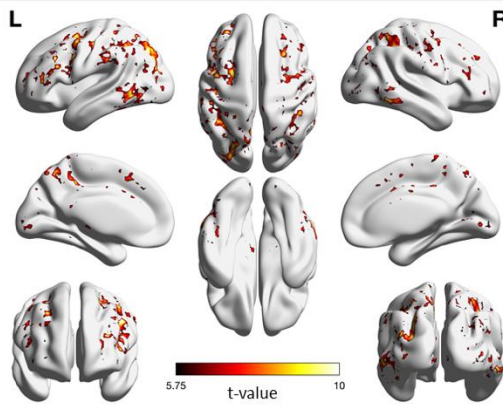
L_FST



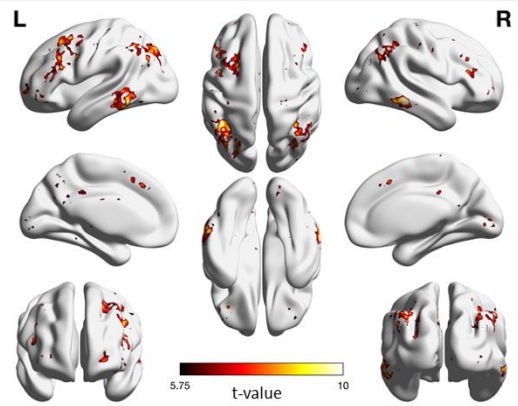
L_PH



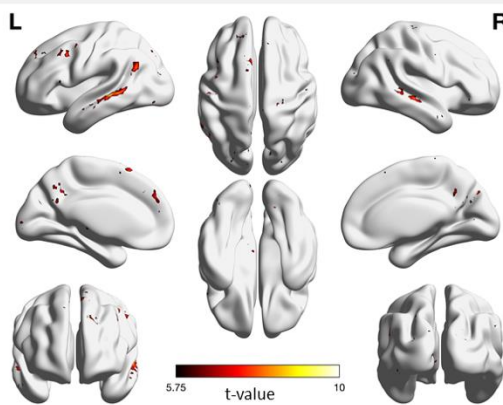
L_PHT



L_TE1p



L_STSvp



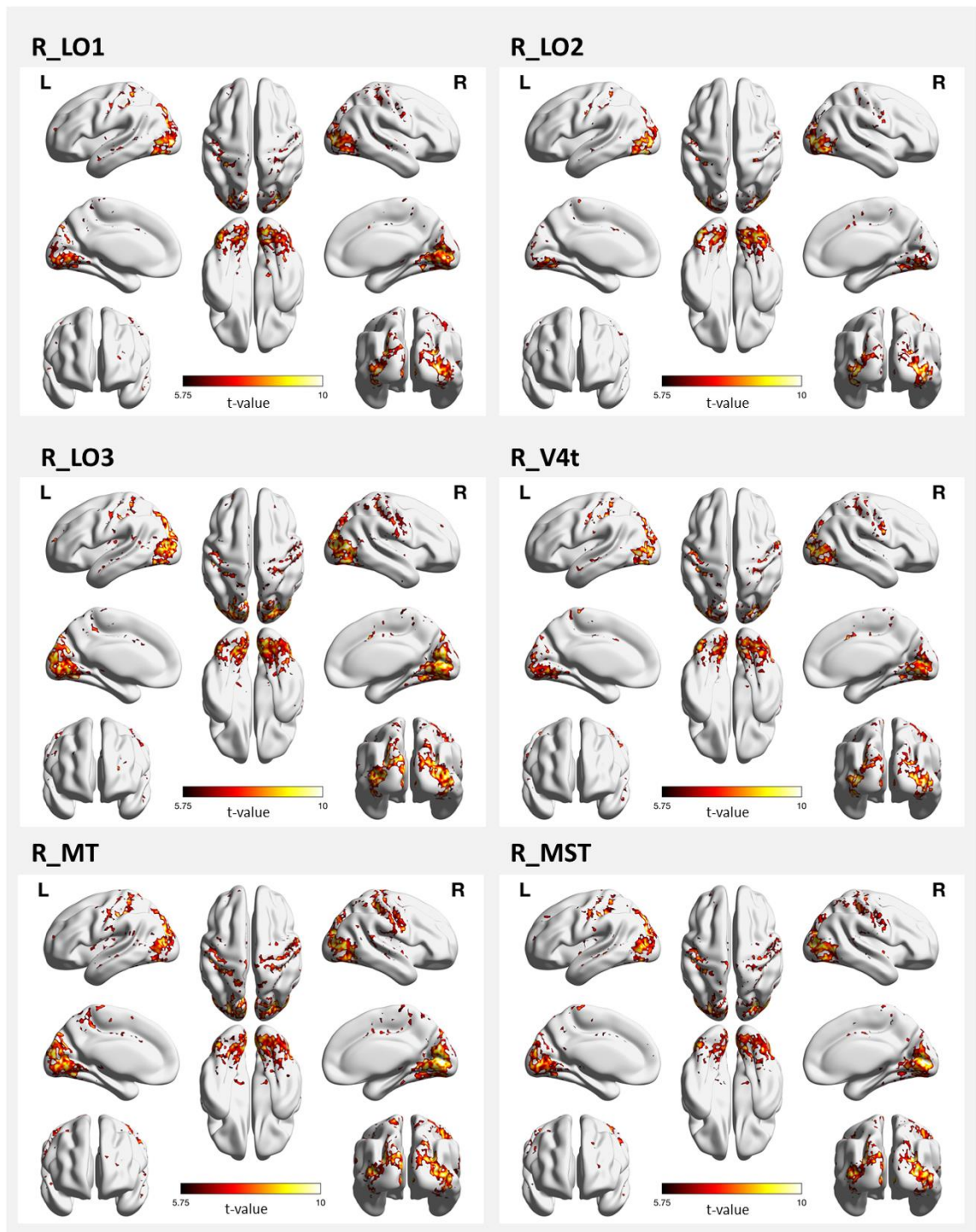
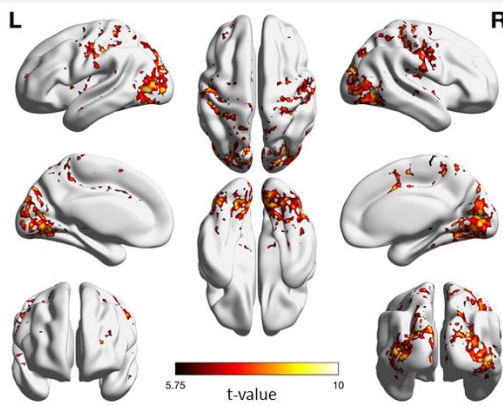
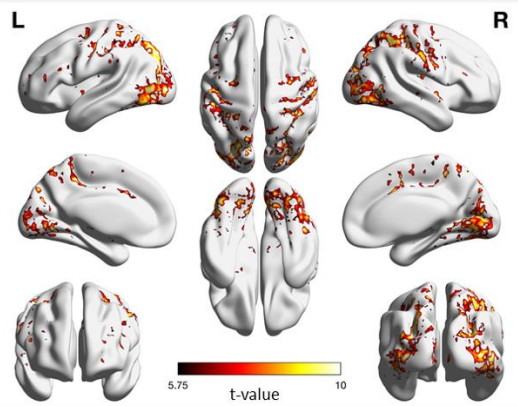


Figure 1 bb

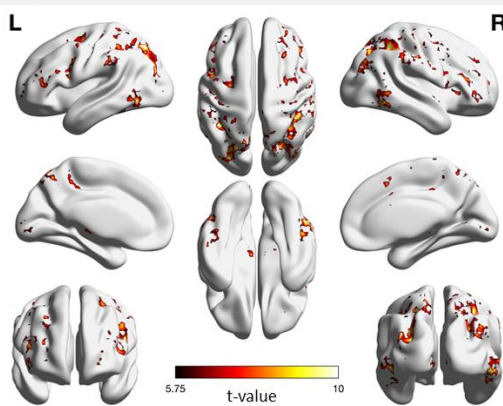
R_FST



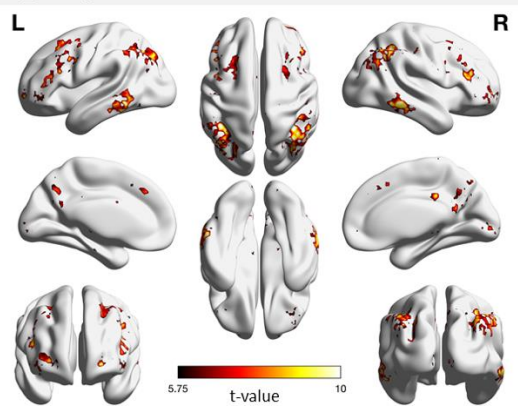
R_PH



R_PHT



R_TE1p



R_STSvp

