

Acknowledgments

The journey to earning this PhD has been like navigating the perilous paths of Middle Earth and the staircases of Hogwarts, full of unexpected challenges, enchanting discoveries and of course, the occasional Balrog and troll under the bridge. There are many people I want to give my gratitude to that have supported me throughout this journey, because this day does not belong to one person but to all, as Aragorn would say.

First, my supervisors. Hans, thank you for believing in me and granting me this opportunity. Thanks for pushing me to make this PhD into the best it could be. Serge, thank you for taking me in in your lab in Amsterdam and supporting this 7T fMRI adventure. Valérie, thank you for being so welcoming and kind.

Thank you to the jury, for reading this thesis and helping me achieve this final step in my PhD journey. Thank you to everyone joining my PhD defense.

Thank you to my interns, Begüm and Roxane, for making this PhD less lonely and helping me out with all this research that had to be done. It was nice to work together.

I would also like to thank all my colleagues in the lab here in Leuven and in Amsterdam. Haemy, for supporting and guiding me already from my internship. The colleagues from the ‘fun office’ for welcoming me in and getting me through the start of this PhD. Carlien and Jelle, thank you so much for being so welcoming, helpful, funny, and kind. You have made my experience in the Netherlands so much better. Gaëlle, we have had many conversations, both serious and fun. You were always there for me both in and outside of work and for that I am so grateful. Thank you for a lovely trip to USA when we went to the SfN conference. Thank you being my friend, I am not sure I would have made it to the end of this PhD without you. Paulien, our friendship is one of the most amazing rewards that this PhD has given me. Right from the start, we just clicked: our humor, our fascination for space, our love for philosophizing about basically everything for hours, and so on. Thanks for all our time spent together and all the support you have given me.

A big thank you to all my friends for providing me with so many positive and fun experiences throughout these years that kept me going! Thanks for boardgame and/or cocktail nights, barbecues, and so on. Thank you, Joycy, for everything! We met during the first year of studying psychology together and we ended up becoming such great friends that have shared so many experiences together already. I’ll be there for you (yes, from the theme song of Friends, let’s sing it at our next karaoke party!). Thank you to all my special partner acrobatics/acroyoga friends! Thank you, Tim, Emily, and Eva for being so welcoming, fun, and kind, you’re all wonderful. Thank you, Hanne and Annabel, I love spending time with you, whether we’re talking, laughing or acro’ing.

Thank you to my family, without you I would not have become the person I am today. Thank you for loving me simply the way I am. Thank you for all the intellectual stimulation that led me to doing and finishing this PhD. Ybe, thanks for accepting me as your weird sister and always being there to play games with me to just relax and have fun. Michiel, thank you for being my big brother and taking care of me throughout our childhood. Bomp and meme, thank you for supporting my intellectual journey and taking care of me, doing this PhD would not have been possible without you. Meme, even though you are not here anymore to witness this, I hope you would be proud. Mama and Peter, thank you for raising me, for providing me with an environment full of humor and openness, for supporting me throughout this crazy life. I love all of you.

At last, I turn to my own little family. I want to thank Mika, my dog, who has been my companion for more than 12 years now. You witnessed all the good and all the bad, you've made me laugh so much and gave me endless cuddles when I did not feel like laughing. Christof, I don't think there are any words that I could say, that would do justice to what you mean to me. There is no one I would rather share this life with than you and I am so grateful to have you by my side. Thank you for all your support throughout this PhD journey: reading texts you did not understand, moving to the Netherlands with me, listening to endless monologues from me, and so on. I love you.

And now finally, mischief managed, PhD completed. And to say it with the words of Bilbo Baggins: I will now step into a new road, and if I don't keep my feet, there is no knowing where I might be swept off to.

Summary

Understanding the extremely diverse and ever-changing world around us is a fundamental skill for all living beings. Through visual categorization, we can recognize objects and shapes such as faces. Visual categorization is supported by a region in the brain termed the occipitotemporal cortex (OTC). This part of the cortex contains multiple areas selective towards certain categories, for example, the fusiform face area and the visual word form area. Compared to the number of categories that we can recognize, only a limited number of category-selective areas have been discovered. The question of how OTC is organized to support such efficient visual categorization, remains unsolved.

In a first methodological study, we compared three of the most popular analyses used in fMRI research to help decide which of these were preferred for our investigation of the organization of OTC. Univariate, representational similarity (also called multivariate) and functional connectivity analysis can all be used to identify functional networks. We discovered that the three analyses result in highly similar networks, but that each network also contains unique information.

In the second study, 7T fMRI provided us with data of a high spatial resolution and a high signal-to-noise ratio for a vast number (20) of visual categories from 19 participants. Guided by the results from the first study, we chose to examine the data using univariate and multivariate analyses. Recent reviews on the visual word form area pressed the need for a study at a high spatial resolution at the level of the individual brain that both explores the various subareas of word selectivity and relates these to other category-selective areas. Thus, with our data, we created a map of category selectivity on the brain surface of each participant and constructed the representational space of OTC to also include the many different dimensions on which categories vary (e.g., animacy). Our research demonstrated among other things, the existence of three subareas of the visual word form area and an additional subarea near the face-selective area that is located in the posterior fusiform gyrus. These findings provide the field with a reliable and consistent method to locate subareas of word selectivity.

In the third study, we focused on the hand and tool areas. This topic could benefit, just like the previous topic about word areas, from the use of 7T fMRI. For example, it could further investigate the possibility that the overlap, found between the hand and tool areas, could have been a result from using a low spatial resolution. In our study, several new aspects compared to previous studies were also added, such as including the faces category. Again, we chose to use univariate and multivariate analyses. Our results showed among other things, that the hand areas differentiated between tools and other objects and that the tool areas distinguished between hands and other categories. Our results provide several confirmations of previous findings, but also add to them in multiple ways, which can inspire new research.

Samenvatting

Het begrijpen van de uiterst diverse en steeds veranderende wereld om ons heen is een fundamentele vaardigheid voor alle levende wezens. Door visuele categorisatie kunnen we objecten en vormen herkennen, zoals gezichten. Visuele categorisatie wordt ondersteund door een regio in de hersenen die de occipitotemporale cortex (OTC) wordt genoemd. Dit deel van de cortex bevat meerdere gebieden die selectief zijn voor bepaalde categorieën, zoals de fusiform face (gezicht) area en de visual word (woord) form area. In vergelijking met het aantal categorieën dat we kunnen herkennen, zijn er slechts een beperkt aantal categorie-selectieve gebieden ontdekt. De vraag rond hoe OTC is georganiseerd om zulke efficiënte visuele categorisatie te ondersteunen, blijft onopgelost.

In een eerste methodologische studie hebben we drie van de meest populaire analyses vergeleken die worden gebruikt in fMRI onderzoek, om te beslissen welke van deze analyses de voorkeur hadden voor ons onderzoek rond de organisatie van OTC. Univariate, representational similarity (ook wel multivariate genoemd) en functional connectivity analyse kunnen allemaal worden gebruikt om functionele netwerken te identificeren. We hebben ontdekt dat de drie analyses zeer vergelijkbare netwerken opleveren, maar dat elk netwerk ook unieke informatie bevat.

In de tweede studie hebben we gebruik gemaakt van 7T fMRI data met een hoge spatiale resolutie en een hoog signaal-ruis ratio voor een groot aantal (20) visuele categorieën van 19 deelnemers. Geleid door de resultaten van de eerste studie, kozen we ervoor om de data te onderzoeken met univariate en multivariate analyses. Recente review artikels van de visual word form area benadrukten de noodzaak van een studie met een hoge spatiale resolutie op het niveau van het individuele brein, die de verschillende subgebieden van woordselectiviteit onderzoekt en deze ook in verband brengt met andere categorie-selectieve gebieden. Zo hebben we met onze data een kaart van categorie selectiviteit op het hersenoppervlak van elke participant gemaakt en construeerden we de zogenaamde representational space om ook de vele verschillende dimensies waarop categorieën variëren (zoals animacy: hoe 'levend' iets is) op te nemen. Ons onderzoek toonde onder meer het bestaan aan van drie subregio's van de visual word form area en een extra subregio nabij de gezicht regio dat zich in de posterior fusiform gyrus bevindt. Deze resultaten voorzien het onderzoeksveld van een betrouwbare en consistente methode om subgebieden van woordselectiviteit te lokaliseren.

In de derde studie focusten we ons op de hand- en gereedschapsgebieden. Dit onderwerp zou, net als het vorige onderwerp over woordgebieden, baat kunnen hebben bij het gebruik van 7T fMRI. Bijvoorbeeld, er kan verder onderzoek gedaan worden naar de mogelijkheid dat de overlap die gevonden is tussen de hand- en gereedschapsgebieden een gevolg is van het gebruik van een lage spatiale resolutie. In onze studie werden ook verschillende nieuwe aspecten toegevoegd in vergelijking met eerdere studies, zoals het toevoegen van de gezichten categorie. Opnieuw kozen we ervoor om univariate en multivariate analyses te gebruiken. Onze resultaten toonden onder andere aan dat de handgebieden onderscheid maakten tussen gereedschappen en andere objecten en dat de gereedschapsgebieden onderscheid maakten tussen handen en andere categorieën. Onze resultaten bevestigden eerdere bevindingen, maar vullen deze ook op meerdere manieren aan, wat nieuw onderzoek kan inspireren.

Contents

<u>ACKNOWLEDGMENTS</u>	<u>I</u>
<u>SUMMARY</u>	<u>III</u>
<u>SAMENVATTING</u>	<u>V</u>
<u>CONTENTS</u>	<u>VII</u>
<u>GENERAL INTRODUCTION</u>	<u>1</u>
THE NATURE OF THE HUMAN BRAIN	1
THE VISUAL SYSTEM: FROM THE EYES TO THE VISUAL CORTEX	3
THE ROOTS OF VISUAL CATEGORIZATION	6
THE CASE OF THE VWFA	11
THE CASE OF HAND AND TOOL AREAS	15
EXPERIMENTAL METHODOLOGIES: fMRI	22
REFERENCES	28
<u>CHAPTER 1: A COMPARISON OF FUNCTIONAL NETWORKS DERIVED FROM REPRESENTATIONAL SIMILARITY, FUNCTIONAL CONNECTIVITY, AND UNIVARIATE ANALYSES</u>	<u>43</u>
ABSTRACT	43
INTRODUCTION	44
MATERIALS AND METHODS	49
DATASETS	49
REGIONS OF INTEREST (ROIs)	50
UNIVARIATE ANALYSIS	51
REPRESENTATIONAL SIMILARITY ANALYSIS	52
FUNCTIONAL CONNECTIVITY ANALYSIS	52
SIGNAL-TO-NOISE RATIO MEASUREMENT	53
ANATOMICAL PROXIMITY	54
COMPARING THE ACTIVATION, REPRESENTATION AND CONNECTIVITY NETWORK	54
RESULTS	56

NETWORKS	56
COMPARING NETWORKS	58
SUMMARY	65
DISCUSSION	65
CONCLUSION	67
DATA AVAILABILITY STATEMENT	68
ETHICS STATEMENT	68
AUTHOR CONTRIBUTIONS	68
FUNDING	68
ACKNOWLEDGEMENTS	68
REFERENCES	68

CHAPTER 2: THE POSITION OF VISUAL WORD FORMS IN THE ANATOMICAL AND REPRESENTATIONAL SPACE OF VISUAL CATEGORIES IN OCCIPITOTEMPORAL CORTEX

ABSTRACT	75
INTRODUCTION	76
METHODS	80
SUBJECTS	80
STIMULI	80
EXPERIMENTAL DESIGN	83
(F)MRI ACQUISITION	83
PREPROCESSING	84
DATA ANALYSES	87
RESULTS	93
FUNCTIONAL NEUROANATOMY OF VENTRAL OTC	93
CHARACTER CATEGORIES GROUP TOGETHER IN REPRESENTATIONAL SPACE AND WORDS DO NOT SHOW A CLEAR COMPETITION WITH OTHER RELEVANT CATEGORIES	107
DISCUSSION	112
SUPPLEMENTARY MATERIAL	119
SUPPLEMENTARY RESULTS	129
REFERENCES	131

CHAPTER 3: A 7T FMRI INVESTIGATION OF HAND AND TOOL AREAS IN THE LATERAL AND VENTRAL OCCIPITOTEMPORAL CORTEX

ABSTRACT	141
INTRODUCTION	142
METHODS	144
PARTICIPANTS	144
STIMULI	144
EXPERIMENTAL DESIGN	147
(F)MRI ACQUISITION	148
PREPROCESSING	149
DATA ANALYSES	152
RESULTS	157
THE OVERLAP BETWEEN HAND AND TOOL AREAS	157
THE RESPONSE PROFILE OF HAND AND TOOL AREAS	160
THE REPRESENTATIONAL SPACE OF HAND AND TOOL AREAS	165
DISCUSSION	169
SUPPLEMENTARY MATERIAL	175
REFERENCES	191
<u>GENERAL DISCUSSION</u>	<u>197</u>
OBJECTIVES	197
MAIN FINDINGS	198
COMPARING FUNCTIONAL NETWORKS REVEALED BY THREE DIFFERENT fMRI ANALYSES	198
THE VISUAL WORD FORM AREA	199
THE HAND AND TOOL AREAS	201
THE BIGGER PICTURE	204
STRENGTHS AND LIMITATIONS	207
FUTURE RESEARCH	211
REFERENCES	213
<u>GENERAL CONCLUSION</u>	<u>217</u>

General Introduction

The nature of the human brain

The brain is a vital organ that controls the nervous system, or more generally, the functioning of the body and mind. For example, it regulates our breathing, it enables you to walk, through it you can plan your entire lives, and feel all kinds of emotions that make living such a rich experience. It is located inside the skull, which protects it. The brain contains billions of neurons, a type of cell, and even more of another type of cell called glial cells that support the neurons. These neurons both process and transmit information using electrical signals, also called the action potential, and chemical signals, also called neurotransmitters. Neurons consist of a cell body or the soma, dendrites, an axon, and synapses (Figure 1). The dendrites receive info from other cells and bring it to the cell body, whereas the axon carries signals away from the neuron to other cells. The synapse is where two neurons connect.

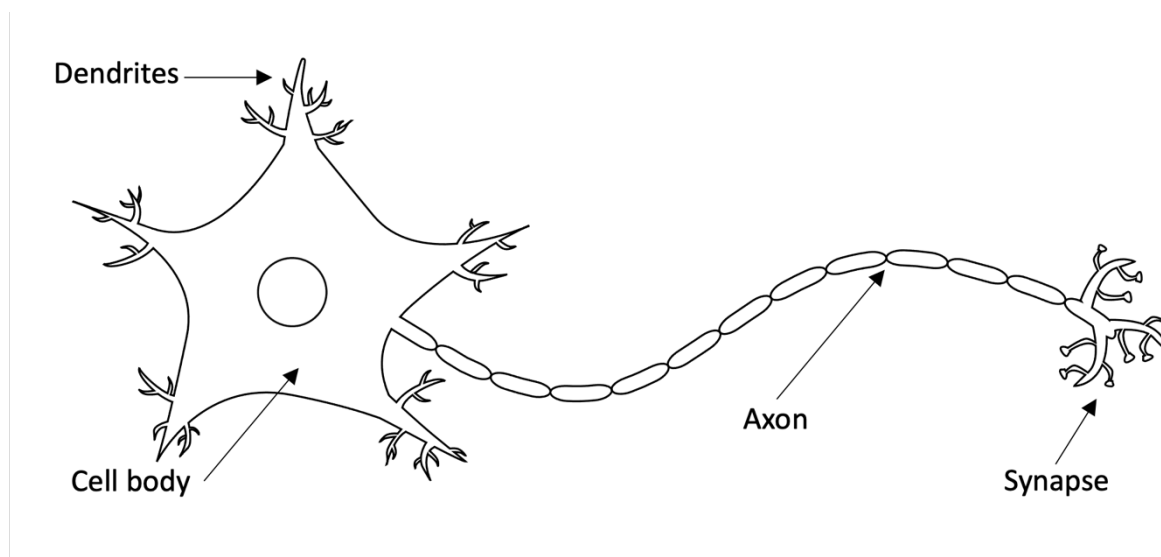


Figure 1. A drawing of neuron and its main parts. Adapted from Zabaleta, 2019: https://commons.wikimedia.org/wiki/File:Drawing_of_a_neuron.svg

The brain itself can be subdivided into several parts (Figure 2).

- (1) The cerebrum: this is the largest part of the brain and enables cognitive functions such as memory, perception, emotions, and so on. The cerebrum can be further subdivided into a left and right part, called hemispheres. These are connected to each other through the corpus callosum, which is located in the center of the cerebrum. This structure is built up of a bundle of nerve fibers, also called white matter tracts, which will be explained later. The cerebrum is not just partitioned into left and right, but also into four different lobes: the occipital,

temporal, parietal and frontal lobes. The occipital lobe allows us to process all the visual information that we collect through our eyes. The temporal lobe supports several different functions such as memory, emotions, and language. The parietal lobe is responsible for processing sensations like touch and taste. The frontal lobe enables our executive functioning like planning and decision-making.

- (2) The cerebellum: this is a much smaller part of the brain. Nonetheless it contains a lot of neurons. It is responsible for coordinated movements. This part of the brain is also subdivided into two hemispheres and into lobes: the anterior, posterior and flocculonodular lobe.
- (3) The brainstem: this structure connects the brain to the spinal cord, which is a part of the nervous system that is responsible for carrying messages to and from the brain. The brainstem is essential for survival as it controls things like our heart rate and breathing. This structure can also be partitioned into the midbrain, pons and medulla oblongata.

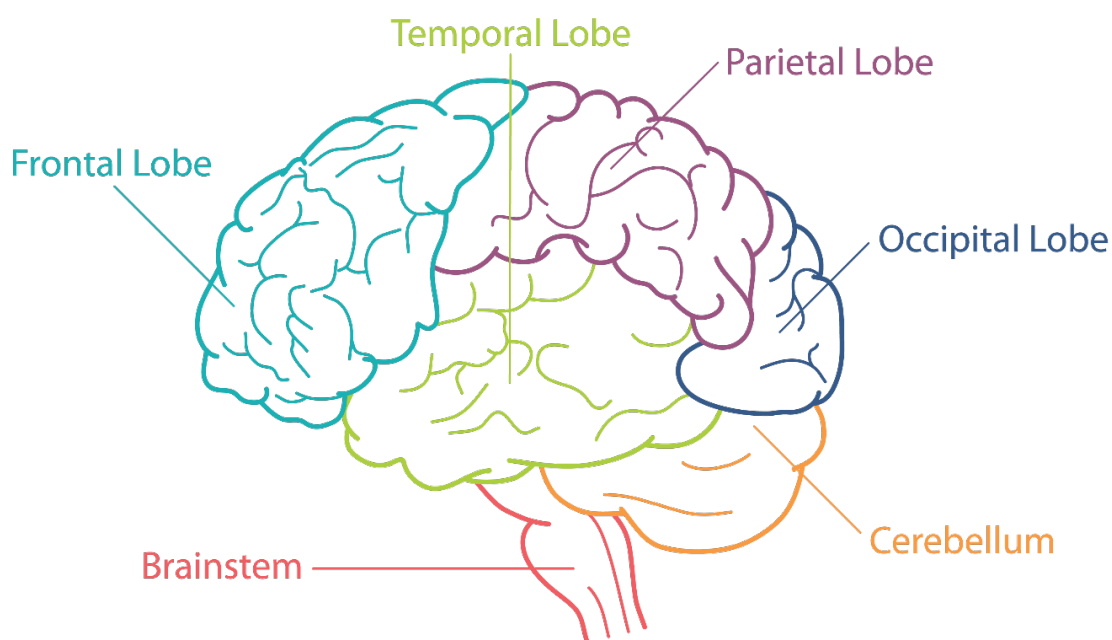


Figure 2. A drawing of the brain and its parts: the cerebrum, consisting of the occipital, temporal, parietal and frontal lobe, and the cerebellum and the brainstem. From https://mhanational.org/sites/default/files/Picture1_0.png

The brain is made up of two types of tissue: gray and white matter (Figure 3). The gray matter is the outer layer of the cerebrum, also called the cerebral cortex. It got its name because this part of the brain looks grayish. It consists of the cell bodies of the neurons. White matter, which is surrounded by the gray matter, consists of the axons of the neurons and thus allow communication between neurons,

close by and far away. These axons are myelinated, which allows fast signal transmission, and they look white. You can imagine the white matter tracts as highways in the brain that are necessary to allow the brain to coordinate between all its parts. Around the white and gray matter, there is cerebrospinal fluid (CSF) that fills all empty spaces within the brain and also the spinal cord. It functions as a sort of cushion between brain and skull that absorbs shocks. It is also responsible for immunological protection and waste regulation.

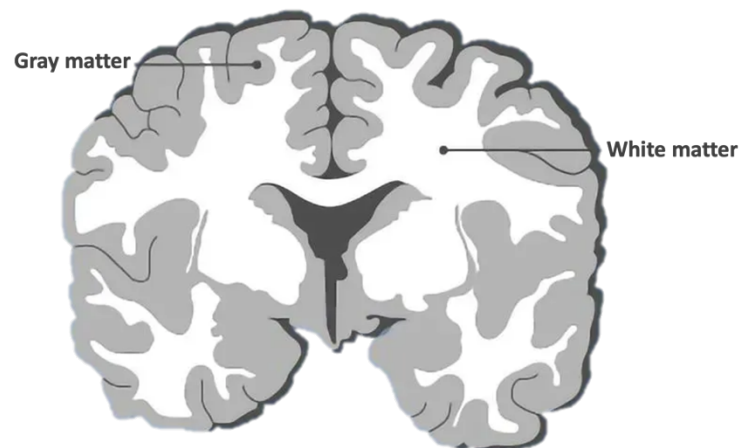


Figure 3. A drawing of a slice of the brain, indicating the gray and white matter. Adapted from <https://assets.technologynetworks.com/production/dynamic/images/content/322973/gray-matter-vs-white-matter-322973-1280x720.webp?cb=12181651>

The visual system: from the eyes to the visual cortex

The visual system allows us to see and interpret what we see. Vision is actually one of the most important senses that we are capable of. About 25% of the brain processes visual information (Van Essen, 2003). Visual processing consists of many stages (Figure 4). The eyes capture light that reflects off objects, through its cornea and its lens. The lens is responsible for projecting the light onto the retina, which is located at the back of the eyes. The retina consists of photoreceptor cells: rods and cones. The rods are responsible for black and white vision and the cones allow color vision. They convert the light into electrical signals through chemical reactions. These signals are transmitted to the brain via the optic nerve. They then reach the optic chiasm. It is here where signals from the nasal part of the retina cross over to the other side. As a result, each hemisphere of the brain processes retinal signals from two eyes, but from only one side of the visual field. The visual information then ends up in a region of the thalamus called the lateral geniculate nucleus (LGN). Finally, through optic radiations,

the info reaches the primary visual cortex (V1), which is located in the occipital lobe of the cerebral cortex.

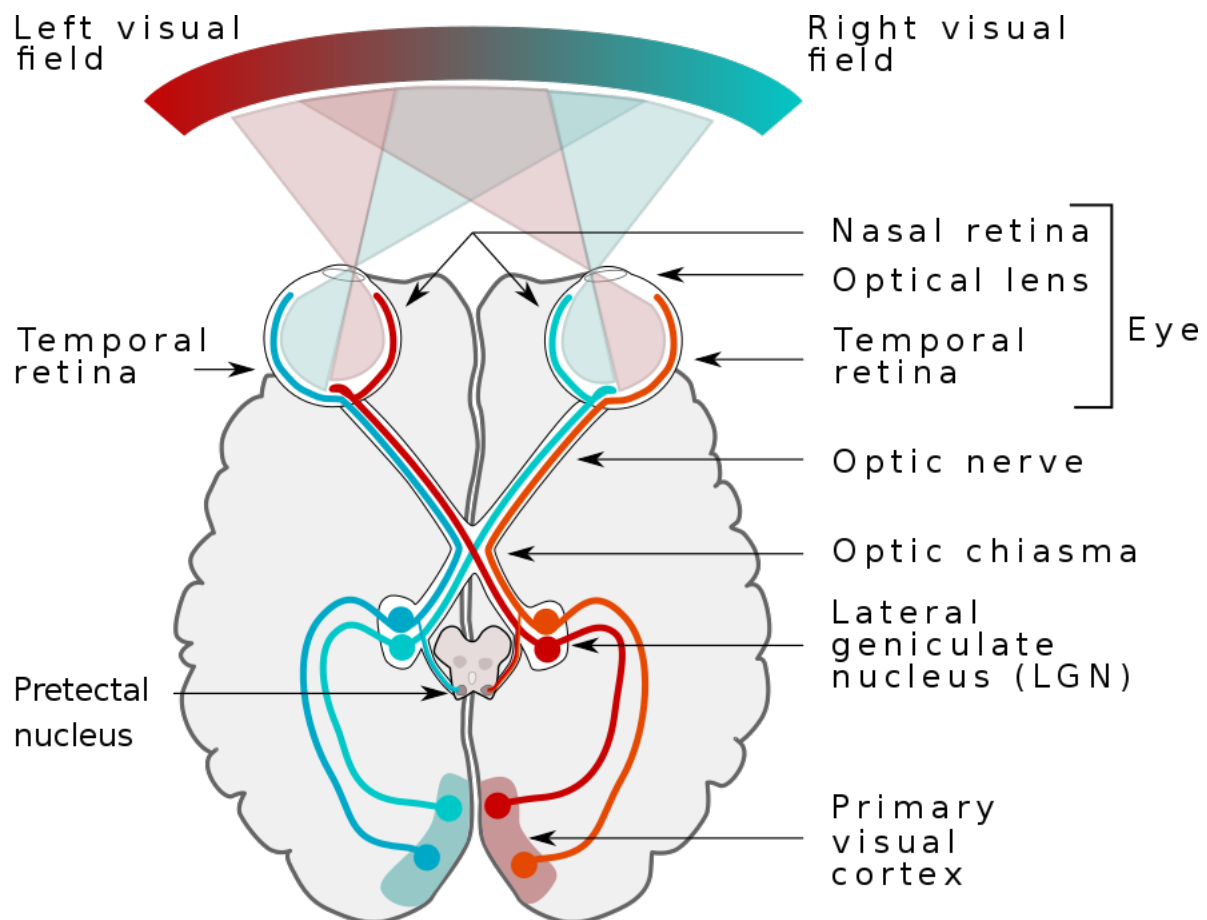


Figure 4. An overview of the stages of the visual system: from the eyes to the primary visual cortex.
 From Nieto, 2015: https://commons.wikimedia.org/wiki/File:Human_visual_pathway.svg

From V1, two types of pathways exist through which the visual info is sent through to other brain areas for further processing (Ungerleider & Mishkin, 1982). The dorsal pathway, or the 'where' and 'how' pathway, is responsible for processing the spatial location of objects and shapes that you are seeing and for processing their movements. It is meant to guide our actions. This dorsal pathway projects from V1 all the way to the posterior parietal cortex. The ventral pathway, or the 'what' pathway, is concerned with recognition of the identity of objects and shapes, and identifying their properties like color and shape. This pathway starts in V1 and projects to the inferotemporal cortex (IT) (Figure 5).

First, in V1, basic or low-level visual features are processed. Examples are the detection of edges, of lines and of orientations, and the processing of contrast and spatial frequency. From this stage to the next, the visual properties that are being processed become increasingly more complex or high-

level. Low-level processing is focused on local parts of the visual image that reached your eye, whereas high-level processing focuses on the whole image (Marr, 1982). The processing thus progresses hierarchically in a feedforward manner, where the results of one stage are passed onto the next stage, and so on (Marr, 1982; Felleman & Van Essen, 1991). This is however not the only way in which information is communicated through the ventral pathway or the brain in general. Processing results from one area are also fed back to previous areas, called feedback, which in turn may influence their processing of information (Gilbert & Li, 2013; Zeki, 2016).

In the second stage, V1 passes info to the secondary visual cortex (V2), where more complex features are processed like texture and size. After this, info passes on to V4 which processes simple geometric shapes and colors. Information ends up in IT, where specific complex shapes are processed, like faces (Kanwisher et al., 1997). Thus, this area of the brain is crucial for object recognition, aka visual categorization. Evidence for this has been provided already long ago (for example, Dean, 1976).

As a final important note, these two pathways are interconnected so that they can provide us with a full picture of our visual environment (Budisavljevic et al., 2018).

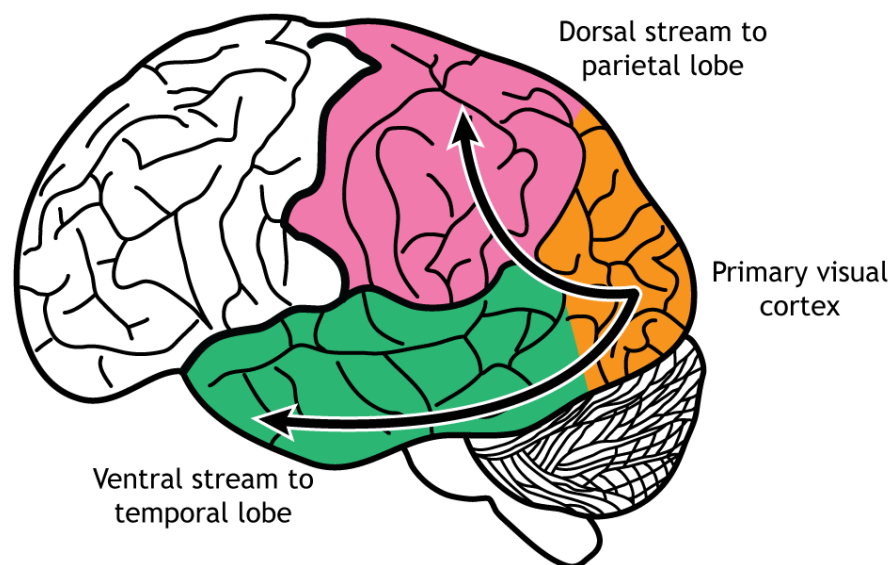


Figure 5. This drawing visualizes the two visual pathways through which visual information is processed. Both start in V1 and progress from there to either the parietal or temporal lobe. From Henley: <https://openbooks.lib.msu.edu/app/uploads/sites/6/2021/03/VisualStreams.png>

Visual categorization is the process by which we can recognize the shapes and objects that we are seeing. By this process, we identify the category of the shape as for example, a face, a chair, and so on. This process is based on the visual properties of the shapes. Thus, it is logical that visual categorization happens at the end of the visual ventral pathway, through which both low and high-level visual features are processed. Through visual categorization, our visual environment makes sense.

Exemplars of a category can vary greatly, and nonetheless we can recognize them as what they are. For example, a chair can have a varying number of legs, different colors and sizes and you can look at it under differing lighting circumstances and viewpoints; and in all these contexts you recognize the object as a chair. And what's more, this recognition is achieved in just an instant, in a tenth of a second (Thorpe et al., 1996; Grill-Spector & Kanwisher, 2005).

The roots of visual categorization

Visual categorization is heavily supported by IT, also often referred to as the occipitotemporal cortex (OTC) of the human brain (Barton et al., 2002; Grill-Spector et al., 2000; Moutoussis & Zeki, 2002; Gaillard et al., 2006). We can investigate this region of the brain with different methods, but in this thesis, we will focus on fMRI, through which we can investigate the activity (patterns) of OTC in response to different conditions (explained in more detail in the Experimental Methodologies section). Through fMRI, already many important observations have been made regarding the functional organization of OTC. A general summary of these can be collected into three main points.

First, the functional organization of OTC consists of multiple areas that respond selectively to specific categories, for example to faces (Kanwisher et al., 1997; Tsao et al., 2006), bodies (Downing et al., 2001; Peelen & Downing, 2005), and scenes (Aguirre et al., 1998; Epstein & Kanwisher, 1998) (Figure 6). These areas are located consistently across participants relative to certain anatomical landmarks and other areas of selectivity (Weiner & Grill-Spector, 2010; Weiner et al., 2014). These studies showed that the mid-fusiform sulcus of the brain predicts the location of one of the face areas present in OTC, and that face areas are organized relative to body areas in a consistent manner.

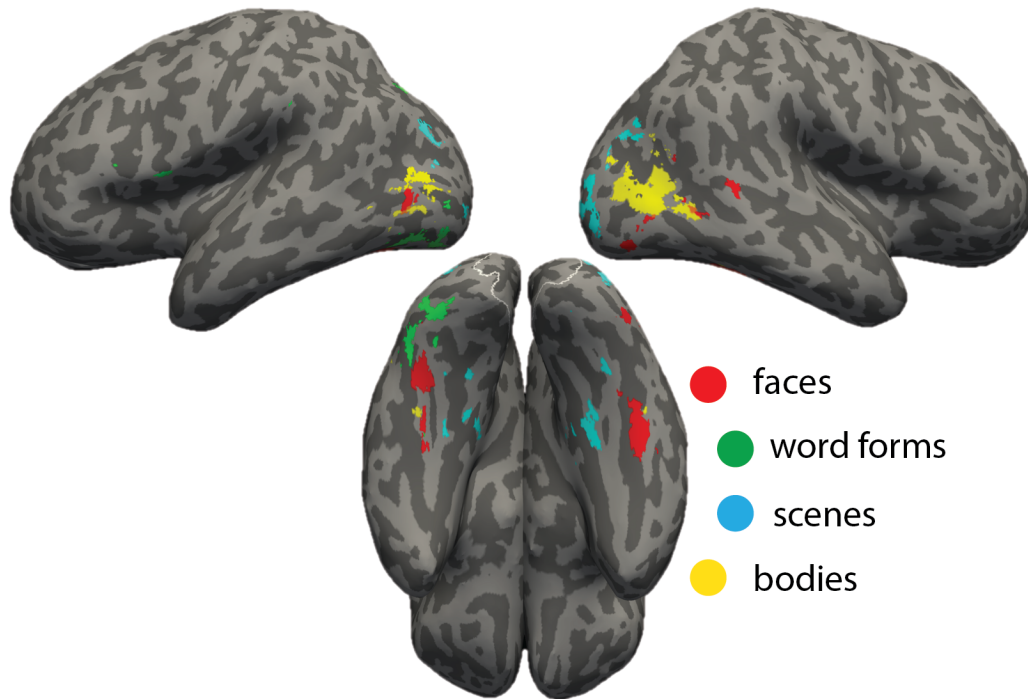


Figure 6. Selective response to four different categories (red: faces, green: word forms, blue: scenes, yellow: bodies) within OTC on an example brain surface. From Op de Beeck, Pillet and Ritchie (2019).

Second, categories, that are not processed by such specific areas within OTC, are represented by different neural activity patterns throughout the whole OTC (Haxby et al., 2001; Kriegeskorte et al., 2008). We call this the representational space of OTC.

Third, the organization of category selectivity within OTC follows specific dimensions on which categories vary. For example, the animacy of a category determines if it is represented more lateral or ventral in OTC: animate categories (e.g., faces) are processed more laterally than inanimate categories (e.g., scenes) (Grill-Spector & Weiner, 2014) (Figure 7, left). This dimension does not only organize where category-selective areas are located within OTC, but also organizes the representational space of OTC (Kriegeskorte et al., 2008). In fact, there are many other organizing dimensions. The real-world size of different categories (Figure 7, right) and body topography is also represented in OTC's functional organization (Orlov et al., 2010; Konkle & Oliva, 2012). Even more low-level dimensions like retinotopy (the location of objects in the visual field) organize OTC (Malach et al., 2002). For example, face areas prefer input that comes from the foveal part of the visual field (Grill-Spector et al., 2017). A recent study proposed a two-dimensional object space that can explain the functional organization of OTC (Bao et al., 2020). They proposed that animacy and aspect ratio are the two main dimensions. However, another recent study has challenged this proposal, concluding that only these two dimensions are not enough to account for OTC's organization (Yargholi & Op de Beeck, 2023).

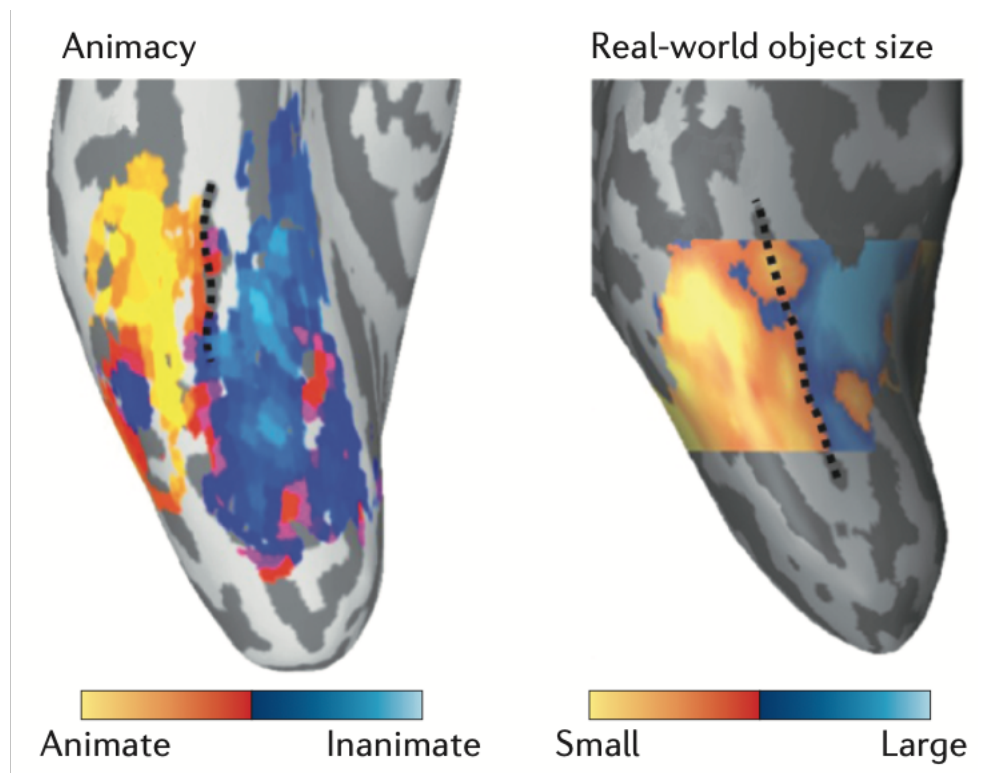


Figure 7. On the left brain, in ventral OTC, animate categories are represented more laterally and inanimate categories more ventrally. On the right brain, again in ventral OTC, objects that are small in real life are represented more laterally and large objects more ventrally. Figure adapted from Grill-Spector and Weiner (2014).

Interestingly, this functional organization of OTC is not limited to humans. It has also been proven that, for example, face areas also exist in primates in IT, which is the proxy of OTC in primates (Tsao et al., 2006). As another example, the animacy dimension that organizes OTC in humans, is also apparent in primates' IT (Kriegeskorte et al., 2008).

There is a long-standing discussion regarding the functional organization of OTC and whether OTC shows true category selectivity, or rather selectivity to visual features that are linked to specific categories. There have been studies that demonstrate how category-selectivity can be explained by low-level features like orientation and spatial frequency (Rice et al., 2014; Watson et al., 2014; Coggan et al., 2016). However, it is a fact that in real life, visual categories are intrinsically linked to visual features and it is therefore expected that OTC would represent categories using a partially feature-based code (Op de Beeck et al., 2019). Thus, the evidence that category-selectivity in part depends on correlated features, does not mean that only these features are being represented (Op de Beeck et al., 2019).

Another issue with the idea that OTC is purely selective to features and not categories, is that studies that have investigated this have used stimuli where features and categories were confounded (Op de Beeck et al., 2019). To provide valid evidence, the two should be separated from each other by creating special stimuli where category information is orthogonal to certain features (Op de Beeck et al., 2019). Several studies have done just this and demonstrated category information remains present in OTC (Bracci & Op de Beeck, 2016; Kaiser et al., 2016; Proklova et al., 2016). To conclude, these findings provide further evidence for the idea that OTC represents categories by using a feature-based categorical code (Op de Beeck et al., 2019).

OTC's role goes beyond categorizing shapes that we know. In fact, research has shown that humans can learn to recognize new shapes. One example study of this was carried out by Op de Beeck and colleagues (2006). They designed several new categories: smoothies, spikies and cubies (Figure 8). After extensive training, participants could categorize these objects. fMRI results showed that the response profile of different areas, including the OTC, of the visual cortex in response to these shapes had changed after training: the response has increased in strength. This demonstrates the brain's plasticity and ability to learn in the case of visual categorization. Many other studies have also provided evidence for this by using several different approaches (e.g., Janzen & Van Turenout, 2004; Weisberg et al., 2007; Wong et al., 2009; Moore et al., 2014; Brants et al., 2016).

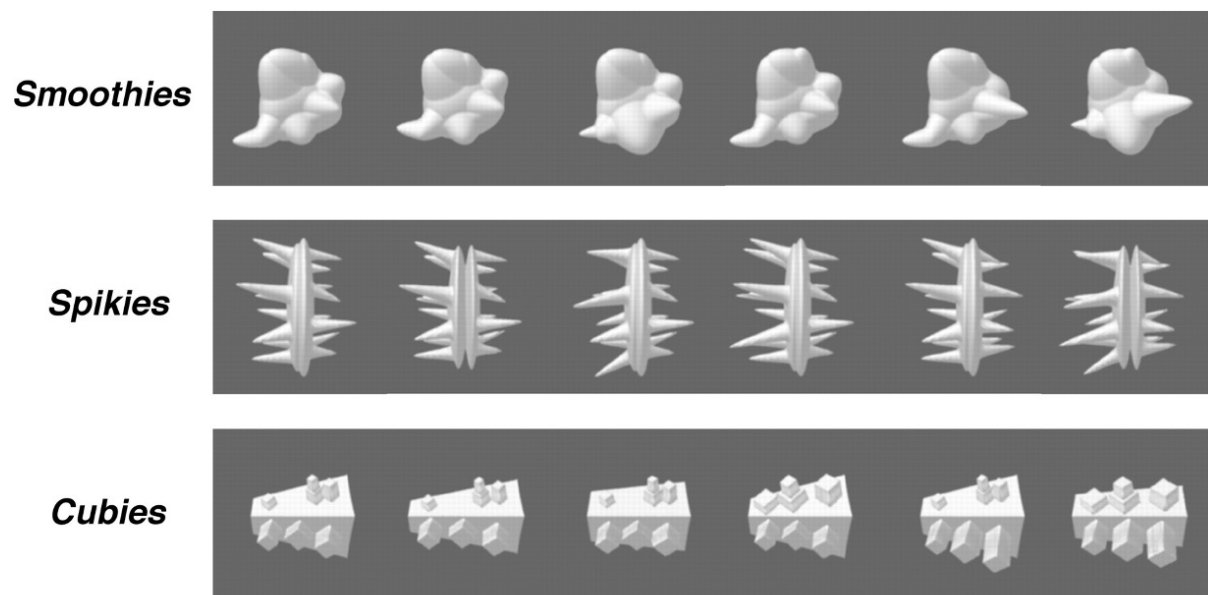


Figure 8. This shows three new types of shapes/objects that participants successfully learned to categorize. Adapted from Op de Beeck and colleagues (2016).

OTC does not only support visual categorization, but categorization in general according to studies with congenitally blind participants. This can appear somewhat unexpected due to OTC's location

inside the visual cortex. One study showed that presenting these people with emotional expressions of voices activates a prominent face area (Fairhall et al., 2017). Another study also investigated the body, object and scene organization of OTC in these participants (also through sounds associated with specific categories) and found this to be similar as compared to sighted participants (van den Hurk et al., 2017). Thus, the OTC in the congenitally blind can be activated through other senses than vision, like audition and touch (Beauchamp, 2005; van den Hurk et al., 2017). Thus, the feature-based categorical code in OTC may in fact be multisensory (Op de Beeck et al., 2019).

One line of studies on OTC concerns itself with how category selectivity develops in this brain region. In Op de Beeck and colleagues (2019), we have proposed a comprehensive framework to explain this. Before we learn to recognize a category, for example faces (starting from a very young age) and words (starting when we learn to read and write in school), several factors are present in the brain that will determine the development of category selectivity (Figure 9). First, the computational hierarchy of the visual system is in place, as described in detail in the visual system section above. Early visual areas process low-level features and later areas can increasingly process more complex features. Second, within the OTC, we propose that organizing dimensions like retinotopy are present. Third, OTC has several paths of connectivity to other areas in the brain, like social and language areas. When we learn categories by experience, these three factors determine where category selective areas will develop. For example, the face areas require high-level computations (holistic processing), selectivity for visual features that are related to faces (retinotopy: faces are often foveated), and connectivity to social areas.

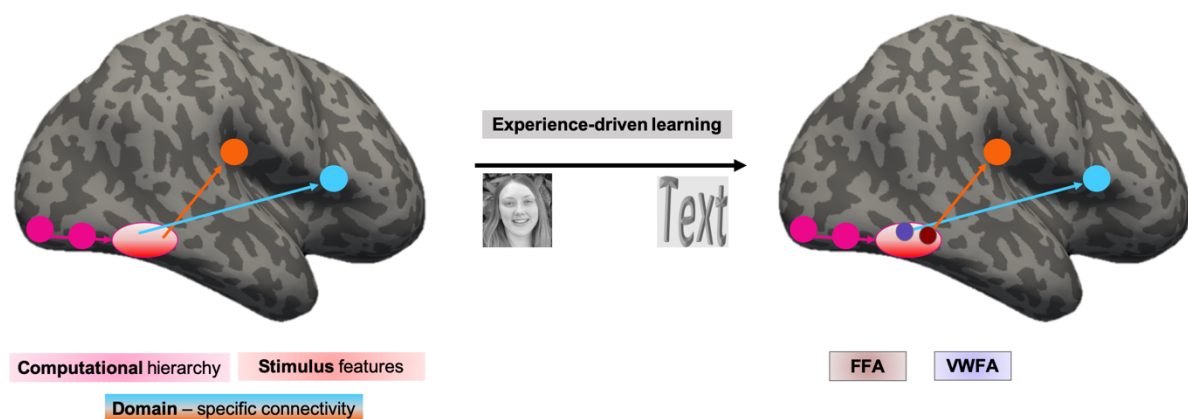


Figure 9. A visualization of the development of category selectivity, constrained by several factors: computational hierarchy (in pink), stimulus features (in red), connectivity specific to the domain (in blue/orange) for two example categories: faces and words. Experiences with and learning these categories leads to the emergence of the fusiform face area (FFA) and the visual word form area (VWFA). Figure from Op de Beeck, Pillet and Ritchie (2019).

On a final note, OTC does not just support visual categorization, it is critical. Several studies have showed how lesions in this brain area lead to various issues with visual categorization depending on where exactly the lesion is and how big it is (Farah, 1990). Konen and colleagues (2011) investigated a patient with agnosia, a name for the impairment of object recognition, that appeared after a brain injury. They demonstrated how the lateral right fusiform gyrus led to this impairment. Schiltz and colleagues (2006) examined a patient with prosopagnosia, an impairment of face processing. They established that the middle fusiform gyrus and inferior occipital gyrus are critical for successful detection and identification of faces.

The case of the VWFA

Reading is one of those remarkable skills that sets our species apart from all other species in the world. Other species have been successfully trained to visually recognize some words, but training has not moved beyond that (Grainger et al., 2012; Srihasam et al., 2012; Rajalingham et al., 2020). As one can imagine, to be able to understand what you are reading, you need to identify all the words at a top speed (like you are doing right now) and this is where visual categorization or object recognition comes into play. Transforming written words into meaning and sounds is achieved in a couple of hundred milliseconds (Sereno & Rayner 2003). Reading starts in the visual cortex but progresses beyond it and needs close communication with other brain areas responsible for linguistic processing and auditory processing (Yeatman et al., 2021).

Learning to read is, evolution-wise, a recent invention and thus, parts of the human brain that originally evolved for other functions, are now recruited to support this reading skill (Yeatman et al., 2021). This is an excellent example of brain plasticity: through hours and hours of experiences that the brain was not really built for per se, the brain adapts itself significantly to allow humans to become complete reading experts.

Learning to recognize faces is different from learning to recognize words. It is believed that face processing uses parts of the brain that have become dedicated to this skill throughout evolution, although this belief is still under debate (Kanwisher, 2010; McKone et al., 2012; Livingstone et al., 2017; Arcaro & Livingstone, 2017; Arcaro et al., 2019). In any case, the brain areas that enable face recognition exist already early in life (Deen et al., 2017) and they develop without explicit training (Yeatman et al., 2021), in contrast to the case of word processing.

Both face and word processing do require category-selective areas within the OTC. In this region of the visual cortex, both face (Kanwisher et al., 1997; Tsao et al., 2006) and word areas (Cohen et al., 2000; Cohen et al., 2002; Baker et al., 2007) exist. Face areas have already been extensively

investigated. For example, how face selectivity can be subdivided into several subareas (including the occipital face area (OFA) and fusiform face area (FFA)) and how these relate to other areas of selectivity (Weiner & Grill-Spector, 2010; Weiner et al., 2014; Grill-Spector et al., 2017).

Word areas, or otherwise called the visual word form area (VWFA), have also received a lot of attention in research. The VWFA extends from the posterior occipitotemporal sulcus (OTS) to the middle of the fusiform gyrus (Yeatman et al., 2021). Areas selective to word forms have a preference for the foveal part of the visual field (Hasson et al., 2002; Rauschecker et al., 2012; Le et al., 2017; Lerma-Usabiaga et al., 2021), similar to face areas (Grill-Spector et al., 2017). Recent studies attempted to subdivide the VWFA, based on computational differences between its subareas, into a VWFA-1 and VWFA-2 (Lerma-Usabiaga et al., 2018; White et al., 2019).

Lerma-Usabiaga and colleagues (2018) contrasted between different conditions to investigate which areas responded to perceptual contrasts and lexical contrasts. They demonstrated that the anterior region (mOTS) was more responsive to lexical features and also more anatomically connected with other language areas, as compared to the posterior part of OTS (pOTS). They believed that the mOTS matches the central VWFA that was previously uncovered described in Vogel et al., 2012; Dehaene et al., 2002; Cohen et al., 2002; Cohen & Dehaene, 2004). According to them the pOTS matches the so-called posterior VWFA from earlier research (described in Vogel et al., 2012; Ben-Shachar et al., 2007). White and colleagues (2019) presented participants with two words at once and then investigated selective spatial attention. They also divided the VWFA in two and found that VWFA-1 (pOTS) could process several words simultaneously. The VWFA-2 (mOTS) processed only one word at a time. Researchers have proposed that VWFA-1 lies next to FFA-1 (also termed pFus) and VWFA-2 next to FFA-2 (mFus), with in between these areas 1 and 2, a body-selective area (Figure 10) (Grill-Spector & Weiner, 2014; Yeatman et al., 2021).

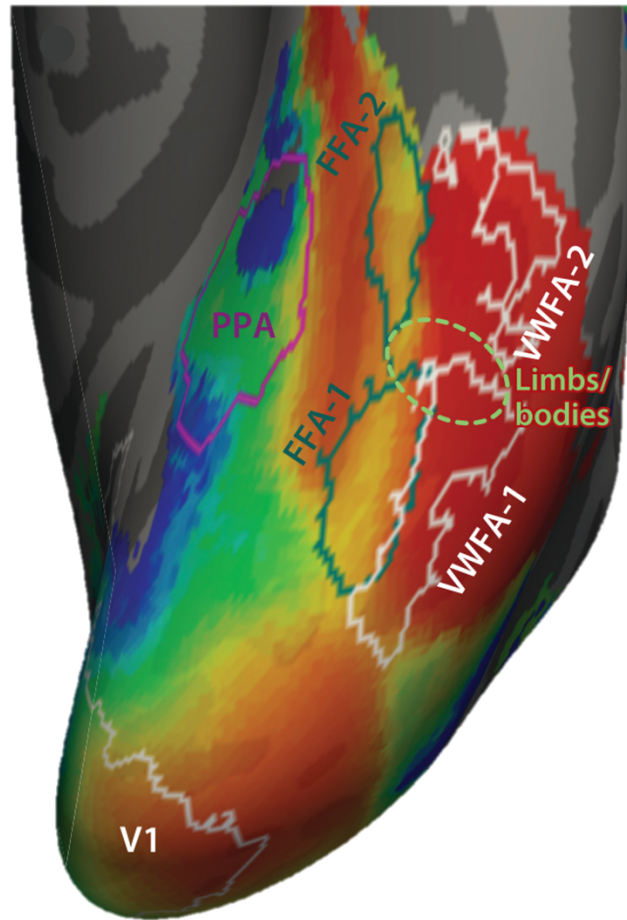


Figure 10. This visualizes the proposal of the layout for VWFA-1 and VWFA-2 in relation to the FFA-1 and FFA-2. Also included are the parahippocampal place area (PPA) and V1. Adapted from Yeatman and colleagues (2021).

It has also been proposed that word areas might follow a processing hierarchy (Yeatman et al., 2021). The idea of hierarchical processing has already been introduced in the section on the visual system, related to the ventral visual pathway. This idea can also be applied to the OTC more specifically given the existence of multiple category-selective areas here that are arranged along a posterior-anterior axis (Yeatman et al., 2021). These observations suggest that there could be hierarchical processing within a category along these areas, or more generally, that these separate areas each support different computations that interact to provide the full picture (Bao et al., 2020). This may be the case for the subareas of face selectivity as well as for subareas of word selectivity. This hypothesis is further supported by the differences in anatomical connections between the VWFA-1 and VWFA-2 (Lerma-Usabiaga et al., 2018).

Interestingly, a line of studies has focused on a right hemifield advantage of visual word recognition and how that relates to the brain. More specifically, this advantage refers to several observations that were made. For example, participants could report more words to the right than to the left of a fixation point and words in the right visual field were recognized better (Huey, 1900; Mishkin & Gorgays, 1952). Another example is that the visual span during crowded letter recognition also extends more to the right than to the left (Legge et al., 2001). And yet another example is that people fixate on a point on the left of the center of a word (so that more letters are on the right) (Rayner, 1998). This advantage is specific to letters and words (Yeatman et al., 2021) and also exists for right-to-left scripts (Almabruk et al., 2011; Siéroff & Haehnel-Benoliel, 2015). The question then is how this advantage is supported by the brain. The belief is that the left hemisphere is the language dominant part of the brain and this is where input from the right hemifield ends up (Yeatman et al., 2021). This hypothesis also matches findings that the VWFA is, like language areas, lateralized to the left hemisphere (van der Haegen et al., 2012; Gerrits et al. 2019) and that the VWFA contains an enhanced representation of the right visual field (Le et al., 2017; White et al., 2019).

A large area of research on the VWFA has concerned itself with its development. Dehaene & Cohen (2007) proposed the neural recycling hypothesis to explain why the VWFA always appears in an anatomically consistent area of the brain across participants. More specifically, they proposed that evolutionary new skills are embedded in parts of the brain that support similar functions. When children learn to read, a competition for cortical territory is then started with face areas and processing (Dehaene & Cohen, 2007), which requires similar functions like analyzing input from the fovea (Behrmann & Plaut, 2020). In the end, face selectivity becomes more lateralized to the right hemisphere while word selectivity takes up territory in the left hemisphere (why the left hemisphere is explained in the previous paragraph) (Dehaene & Cohen, 2011; Behrmann & Plaut, 2020). The part of the hypothesis that posits a destructive competition for territory (Dehaene et al., 2010) is especially under debate. Recent studies showed that the VWFA developed in a region of OTC that was not selective for any category (Saygin et al., 2016; Dehaene-Lambertz et al., 2018) or was selective for many different categories (Kubota et al. 2019). Learning to read may not cause any competition and can even mean a positive change: a fine-tuning of visual object recognition may occur (Hervais-Adelman et al., 2019). In contrast, another recent study by Nordt and colleagues (2021) found that selectivity to faces and words strengthens but selectivity to limbs weakens, which indicates both a fine-tuning and a destructive competition.

The VWFA's role goes beyond visual reading. People that are congenitally blind cannot read using their eyes. They can learn to read Braille, which consists of feeling dotter patterns with their fingers. Around the VWFA's standard location, selectivity to written words in Braille was found in these

participants in several studies (Reich et al., 2011; Kim et al., 2017; Bedny, 2017). And it even goes beyond touch. For example, Striem-Amit and colleagues (2012) showed activation of the VWFA in congenitally blind participants during listening to letter soundscapes. Evidence for the multisensory nature of the VWFA does not just come from the congenitally blind. Mattioni and colleagues (2020) showed word selectivity in response to sounds in the ventral OTC also in sighted participants.

On a final note, issues with reading occur when there are problems with the VWFA. A well-known deficit is dyslexia. Symptoms of this deficit are various, from trouble with reading, to issues with spoken language, as well as problems with controlling spatial attention (White et al., 2019b). Therefore, it is unlikely that it occurs due to one single cause and is instead caused by a combination of risk factors (Yeatman et al., 2021). The left ventral OTC, which includes the VWFA, is often implicated as a location of abnormal activity in people with dyslexia (Richlan et al., 2011; Wandell et al., 2012). In these children, it is sometimes even impossible to identify the VWFA (Kubota et al., 2019). Also the anatomical connections with other areas differ (Boets et al., 2013). With more knowledge and evidence, there is hope for successful interventions though, given the level of plasticity of the brain (Yeatman et al., 2021).

Although there have been many studies investigating different characteristics of the VWFA, some gaps in the literature remain. Recently, prominent researchers in the field have noted the need for studies that investigate the VWFA and its subareas with high anatomical precision on the level of the individual brain (Caffarra et al., 2021; Yeatman et al., 2021). In addition, the location of the VWFA should still be related to other prominent areas of category-selectivity, adding other categories to the mix than just faces. These issues are elucidated in more detail in the introduction of chapter 2 of this thesis, where an investigation is then conducted to remedy these issues.

The case of hand and tool areas

Another interesting case regarding category selectivity in the brain is the existence of hand and tool areas and the overlap between them. Before the identification of a hand area by Bracci and colleagues (2010), a related category-selective area was already identified: the extrastriate body area (EBA) and fusiform body area (FBA), which were found to be right-lateralized, just like face areas (Puce et al., 1996; Kanwisher et al., 1997; Downing et al., 2001; Grill-Spector & Malach, 2004; Peelen & Downing, 2005), although this lateralization is partially dependent on handedness (Willems et al., 2010).

Bracci and colleagues (2010) discovered a hand area in the lateral OTC (lateral occipital sulcus) and proved that this area was separate from the EBA that lay close by. According to this initial study, this hand area was strongly left lateralized, but this was disproved in later research (Bracci et al., 2016).

The authors also noted that the location of this hand area appeared close to where, in other studies, selectivity to tools was found (Chao et al., 1999; Valyear et al., 2007; Bracci et al., 2010). Later studies also identified hand areas in the lateral right OTC and in left and right ventral OTC (Bracci et al., 2012; Bracci et al., 2016). In the studies by Bracci and colleagues in 2010 and in 2012 they contrasted hands against chairs (2010, 2012) or animals (2012), and in 2016 they contrasted with scrambled objects. This may be an explanation as to why in 2012, they also identified a right lateral hand area and then in 2016, they also found ventral hand areas. Selectivity to hands is also found in other parts of the brain than OTC, for example in parietal regions (Bracci & Peelen, 2013; Bracci et al., 2016).

Tools are handheld (often by the dominant hand) objects that are used to achieve certain goals and serve as an extension of our body. Selectivity to tools in the brain is believed to be left lateralized (Mahon et al., 2007; Garcea & Mahon, 2014). As mentioned above, selectivity to tools was found in the left lateral OTC (Chao et al., 1999) but later also in the right OTC (Bracci et al., 2016), in ventral OTC around the medial fusiform gyrus (Chao et al., 1999; Mahon et al., 2007; Garcea & Mahon, 2014; Bracci et al., 2016). Like hand selectivity, tool selectivity is not limited to the OTC and is found also for example in frontoparietal regions (Chao & Martin, 2000; Mahon et al., 2013). Garcea and Mahon (2014) proposed that the tool areas in ventral OTC represent visual, semantic, and surface texture information and lateral OTC tool areas represent movements associated with tool use.

Studies investigating both hand and tool areas showed that part of the hand and tool selectivity in lateral OTC overlapped with each other (e.g., Bracci et al., 2012; Bracci & Peelen, 2013). Bracci and colleagues (2012) confirmed that the hand and tool areas differed from nearby object and motion areas and that the overlap between them was specific to these two categories (much less overlap between tools and bodies). Later research showed the overlap also appeared in right lateral OTC and left and right ventral OTC, and not just in left lateral OTC, and it even showed up beyond the OTC in the intraparietal sulcus (IPS) (Figure 11) (Bracci et al., 2016).

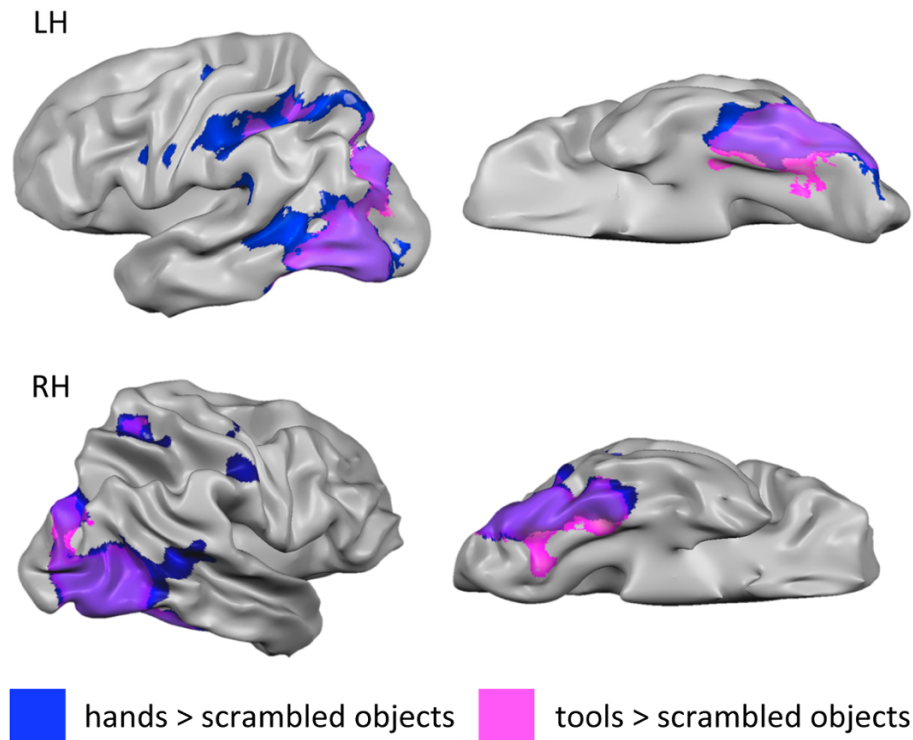


Figure 11. In blue selectivity to hands, in pink selectivity to tools, and in purple the overlapping activity between hands and tools, in the left hemisphere (top left: lateral OTC, top right: ventral OTC) and in the right hemisphere (bottom left: lateral OTC, bottom right: ventral OTC). Adapted from Bracci and colleagues (2016).

The overlap between hand and tool areas in OTC, part of the visual cortex, was unexpected, as they do not look visually alike. As explained earlier, the organization of OTC follows several category dimensions, for example animacy. Also with these dimensions, the overlap cannot be explained (as e.g., hands are considered animate and tools inanimate). A line of research is thus dedicated to understanding the cause of this overlap. One line of research looks into less obvious visual features as an explanation, whereas another line looks into shared information related to the action domain as the cause.

First, several vision-based theories exist. The small size of both might be a reason, as real-world size organizes OTC (Konkle & Oliva, 2012), or the elongated shape of tools relative to other categories used in a contrast to define tool selectivity (Sakuraba et al., 2012). Freud and colleagues (2017) investigated factors such as these. They created tools and non-tool objects using a 3D printer, matched on several variables: elongation, size, viewing distance, material properties and graspability. During fMRI, participants viewed images of these. Results showed a higher activation to tools compared to

non-tools in lateral OTC, showing that, although no hands stimuli were included, tool selectivity in OTC cannot be explained by such low-level visual features. Evidence from congenitally blind show a similarly organized OTC, which implies visual features (alone) cannot account for the way category selectivity in OTC is organized (Pietrini et al., 2004; Amedi et al., 2007; Sathian & Lacey, 2007; Mahon et al., 2009; Reich et al., 2011; Striem-Amit et al., 2012; Peelen et al., 2013; Kim et al., 2017; Bedny, 2017; Mattionni et al., 2020). Further evidence by Perini and colleagues (2014) showed that during transcranial magnetic stimulation (TMS) of the hand and tool overlap in lateral OTC, participants viewing tool images were impaired at an action knowledge task but not the typical location knowledge task. If a visual feature would explain the overlap, TMS would affect performance of identical images in both conceptual tasks in the same way (Perini et al., 2014).

Another visual account concerns the higher co-occurrence of tools and hands in life, compared to other objects and hands. This theory suggests that seeing either hands or tools will activate the imagination of the other category. This is somewhat related to Orlov and colleagues (2010) who proposed that face and body areas lie close together because they are seen together often in real life, and this idea may also be applied to hands and tools. Several things speak against this mental imagery account. First, O'Craven and Kanwisher (2000) showed that activation is greater for directly viewing an image as compared to mental imagery of the stimulus. Thus, if this account would explain the overlap, we would observe a weaker response for the indirectly activated category and this is not the case (Bracci et al., 2010, 2012; Bracci & Peelen, 2013; Peeters et al., 2013). Second, areas activated through mental imagery of a category should be limited to areas active when directly viewing the relevant category, and this was not observed as the hand and tool area only partially overlapped (Bracci et al., 2012). Third, arms and hands for example also co-occur like tools and hands, but for this pair the overlap was significantly smaller than for hands and tools (Bracci et al., 2012). Based on the previously introduced TMS study, based on the visual imagery account, results would have had to show that TMS did not influence performance on either the action knowledge task or the typical location knowledge task, as the overlap area would not serve a role critical to either of these tasks (Perini et al., 2014). And once again, the evidence from congenitally blind also strongly speaks against the visual imagery account (Pietrini et al., 2004; Amedi et al., 2007; Sathian & Lacey, 2007; Mahon et al., 2009; Reich et al., 2011; Striem-Amit et al., 2012; Peelen et al., 2013; Kim et al., 2017; Bedny, 2017; Mattionni et al., 2020). As a result of all these findings, one may believe that OTC did not develop just for (visual) perception, but for information extraction (Cisek, 2007).

Second, a shared feature between hands and tools from the action domain might explain the overlap. A first idea comes from Magri and colleagues (2016) who showed that lateral OTC preferred interactive objects above less interactive ones, which might imply that motor execution of object

categories plays a role. However, Tarhan and colleagues (2015) showed that lesions in lateral OTC resulted in impairments of recognition of tool related actions, but not of execution of these actions. Also, people born without hands show hand and tool selectivity (Striem-Amit et al., 2017). Research on prosthesis users is also not consistent with this account (van den Heiligenberg et al., 2015; van den Heiligenberg et al., 2018; will be explained in more detail later).

A second idea concerns manipulability/graspability of an object. Merely seeing a tool could evoke the representation of the action that could be used on that object, that also involves the hands for which the representation is thus also activated (Grèzes et al., 2003). Studies have contrasted graspable and non-tool graspable objects and found that the response to tools was higher than what activation to graspable objects alone could account for (Valyear et al., 2007; Bracci & Peelen, 2013). This is evidence against the graspability explanation for the overlap between hand and tool areas. Knights and colleagues (2021) go further with this graspability idea and investigate if lateral OTC encoded how tools should be grasped, as proposed by (Mahon & Caramazza, 2009; Striem-Amit et al., 2017). They asked participants to grasp tool and non-tool objects (matched for elongation, width and depth), that were 3D printed, in a(n) (a)typical way. The participants were not asked to use the tool. Whether or not the tool was grasped typically, was expected to influence activation in the left lateral hand area, only for tools and not for the non-tools. Results showed exactly that. This effect was specific to the left lateral hand area though, as the same results were not found in tool areas, object areas or body areas, even if they overlapped the left lateral hand area and may thus not explain the overlap between hand and tool areas. Nonetheless, these findings propose an interesting new direction to investigate further.

The third idea relates to hand-action centrality, which reflects the degree that a hand action is a central characteristic of the object (Mahon et al., 2007). Mechanical tools, sports objects and everyday objects like pens all score high on this dimension and all these are action effectors (Bracci & Peelen, 2013). The research also included objects that score high on the dimension but are not action effectors. They found stronger responses to hand-action central effector objects (which includes tools) than to hand-action central non-effector objects. Thus, hand-action centrality cannot explain the overlap between hand and tool areas found in OTC and instead points to a dimension that can be described as objects with which you perform an action. Related to the hand-action centrality idea, is causality: tools have a causal effect on the environment. Tool areas were found to respond more to causer (of the next event) objects than reactor (result of previous event) objects (Leshinskaya et al., 2020). However, Bracci and Peelen (2013) showed that causality alone cannot explain the overlap (see next paragraph for more details).

A fourth proposal refers to body extension, which is defined as the degree to which the body controls an object to interact with the world, in such a manner that the object physically and

functionally extends the boundaries of the body. For example, you could paint with your hand, or with a paint roller which extends the hand physically and needs to be controlled by the body to paint successfully. As such, using tools, a range of goals can be achieved that are not possible with our bodies alone (Perini et al., 2014). Body effectors like hands and object effectors like tools that extend the body, are then expected to be clustered together. This account resembles the causality account in that objects that you act with are similar to causers, and objects acted upon are similar to reactors, but they are different. A light switch for example does not rate high on body extension but does rate high on causality. The body extension account thus refers to causers that also extend the body. Bracci and Peelen (2013) found that the hand area in lateral OTC responded more to objects used as action effectors that also scored high on the body extension account.

Bracci and colleagues (2012) investigated the functional connectivity of the overlap between on the one hand, the hand and tool area overlap in lateral OTC, and on the other hand, nearby body, object, and motion areas, with other areas in the brain. They found that this hand and tool overlap region showed selective connectivity to left anterior intraparietal sulcus (aIPS) and the dorsal premotor cortex (dPM), brain areas that have been shown to be active during observation and execution of hand and tool action (Bach et al., 2010; Valyear & Culham 2010). Furthermore, Amaral and colleagues (2021) found that this overlap area was differentially functionally connected to other areas and networks in the brain depending on the stimulus (tools or hands) that was being processed.

The evidence of these studies is consistent with the hypothesis that connectivity demands can determine where areas develop (see Op de Beeck et al., 2019 for a comprehensive framework). Other studies have also provided evidence for functional connectivity between OTC areas and relevant domain-specific networks (Mahon et al., 2007; Bracci et al., 2016). In the case of the hand and tool areas, body info related to actions, proposedly supported by the OTC, is relevant for action networks and thus areas that process this should develop in a location of OTC where connectivity to these action networks is established (Peelen & Caramazza, 2010; Mahon & Caramazza, 2011). Also, the possible left lateralization of hand and tool overlap might thus be due to the necessity for connectivity to left-lateralized action networks, as the left hemisphere is specialized in visual and motor control (Serrien et al., 2006). This reasoning is similar for the previously discussed VWFA, which is supposedly left lateralized due to strong connectivity in the left hemisphere with language areas (Dehaene et al., 2005; Gaillard et al., 2006), which can in turn explain the right hemifield advantage for reading (Yeatman et al., 2021).

Connectivity demands may also explain findings from another line of studies regarding hand and tool areas and category-selective areas in OTC in general, in the congenitally blind and dysplasics (people born without a hand). Research has shown that, beside the lack of visual or motor experiences,

the organization of category selectivity follows the organization of the healthy population (Pietrini et al., 2004; Amedi et al., 2007; Sathian & Lacey, 2007; Mahon et al., 2009; Reich et al., 2011; Peelen et al., 2013; Striem-Amit et al., 2017; Mattioni et al., 2020). This does not necessarily mean absolutely no experience is required for the development of category selectivity in OTC, but rather that experience in one of the modalities may be sufficient. This also implies that the nature of at least some category-selective areas in OTC is multisensory (Op de Beeck et al., 2019).

Another study related to the influence of experience on category selectivity in OTC, is the study by Schone and colleagues (2021). They investigated the hand and tool selectivity in conjunction in expert tool users (litter pickers) and under the theory of embodiment, expected these participants to embody the tool compared to novice tool users. Specifically, they hypothesized that the representation for tools would be more alike the representation for hands in expert tool users. Results were actually opposite: the representation for the expert and non-expert tools was more dissimilar to that of hands in the expert tool users, as compared to the novice tool users. This is evidence against the embodiment theory and instead, neural differentiation due to experience was demonstrated.

This is not unlike other previous evidence. For example, Op de Beeck and Baker (2010) demonstrated that visual tool experience shaped representations of the visual cortex. It is possible that the representation for the action associated with the tool became more differentiated, instead of the tool on itself (Schone et al., 2021). This could explain why also the representation for the other tool in which they were not experts, became more differentiated, as this non-expert tool was associated with the same type of action (Schone et al., 2021). Another explanation could be that it was not the representation of the tool that changed, but the representation of hands, and this would also explain the differentiation of both expert and non-expert tools in the expert tool users (Schone et al., 2021). These results are also consistent with a study in amputees on prosthetics usage, where the prostheses were more dissociated from hands and tools in the hand-selective portion of EBA (Maimon-Mor & Makin, 2020).

Prostheses are an interesting topic of investigation in relation to the hand and tool areas. There are cosmetic prostheses that look like hands but serve as a tool to the amputee (they elongate the stump), and there are functional prostheses that do not look like hands (a hook) but can be used like hands (e.g., grasping, van den Heiligenberg et al., 2015). In a study on these two types of prostheses, researchers showed there was a correlation between the use of the prosthesis (for both types) and activation in lateral OTC (and also in the IPS), and these areas overlapped with hand and tool areas (van den Heiligenberg et al., 2015). They also found that the correlation with usage could be explained by action knowledge and not visual features: cosmetic prostheses (looking like hands but used as a tool by the amputees), showed a left-lateralized pattern consistent with the organization of tool

selectivity, and functional prostheses (not looking like a hand but used like one) showed a bilateral representation consistent with the organization of hand selectivity (van den Heiligenberg et al., 2015). In a later study, van den Heiligenberg and colleagues (2018) also showed that both type of prosthesis users showed a higher activation in the lateral hand area compared to control participants, even when viewing images of prostheses that were visually different from their own (but still had the same functionality). This proves that the change is not due to familiarity with visual features of the prosthesis, but rather a categorical change in the representations, and this is functional given that prosthesis users routinely replace their prosthesis and use multiple types of prostheses (van den Heiligenberg et al., 2018). They even included controls with passive viewing experience of prostheses (such as family members of amputees), and despite this visual experience, they did not show increased activity in the lateral hand area, as compared to the prosthesis users (van den Heiligenberg et al., 2018). This is further evidence for the theory that the neural changes in prosthesis users are not due to visual experience.

Although a lot of research has been performed on the hand and/or tool areas, and also some on their overlap, several gaps remain. For example, not much is known about the neural dynamics of the response in these areas and this could provide further insight into what could cause the overlap between these areas. Another example concerns the use of 3T fMRI, which limits the spatial resolution at which the brain is investigated. These studies have also used several spatial preprocessing steps that could have influenced the results. These issues are further described, together with an attempt to remedy them, in chapter 3 of this thesis.

Experimental methodologies: fMRI

Magnetic resonance imaging (MRI) is a method used to study the anatomical and functional characteristics of the brain. It is suitable for both healthy and diseased people of most ages and available in most hospitals. It is not an invasive procedure: you lie (still) inside a scanner for a while and either rest or perform a task (for functional MRI or fMRI) and, although it may be a tad uncomfortable (small space, noisy), there are no long-term effects from the machine. The machine operates by strong electromagnets. These manipulate the spinning direction of hydrogen atoms inside the body. Due to these magnets, you cannot bring anything magnetic into the scanner room (like a belt, coins, glasses, ...). Because of this, some participants are not suitable to be scanned, if they have metal in their body, for example a pacemaker. The strength of the magnet is expressed in Tesla (T), and the standard strength used in research nowadays is 3T. Cars can be picked up with a magnet of 1T, so the magnets inside these machines are very strong.

Through the use of fMRI, we can investigate brain activity relative to task conditions included in our experiment. It is sensitive to changes in blood flow. When a brain region becomes active, the blood supply to the region is increased to provide the region with oxygen. Neurons consume oxygen when active and oxygenated blood becomes deoxygenated. More specifically, the ratio of oxygenated and deoxygenated blood is picked up with fMRI and this relates to neural activity, and this is called the blood oxygenation level-dependent (BOLD) signal (Logothetis, 2002).

The brain is imaged as many different voxels (= 3D pixels). The size of the voxels is also called the spatial resolution. Although the spatial resolution (around 2 to 3 mm voxel size with a 3T scanner) is quite high, the temporal resolution is not. A full scan of the brain is taken every couple of seconds, whereas neural processing happens way faster than that (remember that visual categorization happens in the tenth of a second). Through fMRI it is thus hard to conclude anything about neural dynamics.

The brain is always active, no matter the type of task the participant is performing while inside the scanner. To relate activity in the brain specifically to task conditions, it is important to compare activity in response to a condition, to a carefully chosen baseline, that only differs from the condition in the way that is interesting to us. For example, if we want to investigate which areas in the brain are important to process emotional expressions on a face, we might want to contrast face images, with all kinds of emotional expressions, with face images that have been rated as having a neutral emotional expression.

Preprocessing

Once fMRI data is collected (and also an anatomical MRI scan of the brain, that has a higher spatial resolution than fMRI), several steps need to be taken to preprocess the data. These steps are important to carry out before any analyses can take place. A first step would be manual quality control of the images: is there anything weird with the images like a piece of the brain scan missing due to an issue with the scanner. Then slice timing correction may be needed in certain cases. This step corrects for the fact that some pieces of the brain are scanned at another timepoint than the other pieces of the brain (the brain is scanned in pieces/slices across a couple of seconds). After, motion correction is necessary. Even though participants are instructed to lie still, it is impossible to remain perfectly still. Small movements can be modelled and corrected in this step so that throughout the brain scans, an anatomical point (e.g., the corpus callosum) is always in exactly the same spot, so that we can compare scans to each other. Then, if making conclusions across multiple participants' brain scans, the data of every participant needs to be spatially normalized to a standard template like the MNI template (brain template based on 152 anatomical scans). This is because every person's brain differs in shape, and we need to account for that variability before we can include a whole group of people into a group

analysis. Then, since fMRI data also contains a lot of noise signal, the SNR can be increased by spatially smoothing the data (the spatial resolution does decrease this way though). A last step one can take is surface reconstruction based on the anatomy of a person's brain. This allows to visualize the brain activity on a 2D (instead of a 3D) model of the gray matter of the brain where gyral and sulcal patterns of the brain are visible and can thus be related to the found activity.

Univariate analysis

We model the activity (BOLD signal) over time in each brain voxel by using multiple linear regression, also called a general linear model (GLM). In this regression, the BOLD signal over time is the dependent variable, which we are trying to predict with the expected BOLD time series, which we built up based on the variables/conditions of the experiment/task. By this process, we get, for every voxel and condition, a beta value that is an estimation of the neural activity in response to that condition. For each voxel, a statistical contrast can then be made between a condition or conditions of interest versus a baseline/control condition, to find out if the voxel is responsive to the condition of interest. We can do this for every voxel in the brain separately (termed whole-brain analysis), or we can define a region of interest (ROI). This region contains multiple voxels and is defined based on previous research, anatomical characteristics, or a specific hypothesis, etcetera. We can then look at the beta value of a condition averaged across all voxels of the ROI, and then for example, compare conditions or ROIs to each other. This process is visualized in Figure 12. This method is used in all three chapters of this thesis.

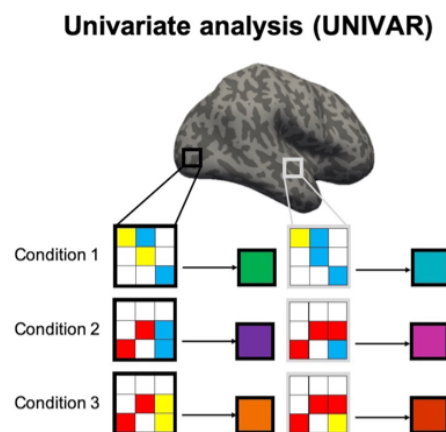


Figure 12. This figure visualizes two brain regions (one with a black outlined and one with a gray outlined square). The smaller squares inside are the voxels and the colors signify the neural activity (the beta values) of the voxels. On different rows, we find the activity of the voxels within the brain regions in response to different conditions. We can then contrast one condition versus another, per voxel; or average across the voxels and then start comparing between conditions and regions of interest. Adapted from Pillet and colleagues, 2019.

Multivariate or multi-voxel pattern analysis (MVPA)

In multivariate analysis, it is not the activity of the voxels that is of interest, but rather the pattern of activity that multiple voxels form together, for one condition compared to another condition. This pattern is also termed the neural representation. It is believed that this pattern may contain information about the condition. This method was first proposed by Haxby and colleagues (2001), who correlated the pattern for one condition with another and could then, for example, investigate if the patterns of two given conditions were more similar than another pair of conditions. This method is called correlational MVPA (Figure 13). One can save the results in a so-called representational dissimilarity matrix (RDM). Each row and column in such a matrix stand for a certain condition and thus in each cell, there is the correlation between the neural pattern (on average) of one condition with another condition. While the average activation of two regions, based on univariate analysis, may not differ, the neural pattern of the two regions may very well be different. MVPA is thus another way of investigating the role of a group of voxels or an ROI in the processing of experimental conditions. This MVPA method is used in chapter 1 of this thesis.

Another method is to train a so-called classifier, a sort of computer algorithm, to predict which condition any given neural pattern is. The classifier is trained with part of the data and then tested with another part of the data. We can then look at the accuracy of the classifier. If the classifier had to guess between 'this is a face' or 'this is a house', we would expect that the neural patterns of these conditions are very different and thus the classifier could do a good job at predicting. However, if it had to guess between 'this is a face with a large nose' or 'this is a face with a slightly less large nose', the neural patterns would probably not differ much for these two types of images and thus the classifier could not have learned to predict this well (and would thus show a bad accuracy). So, the higher the accuracy, the more dissimilar/distinguishable the patterns of two conditions were. The (average) accuracy of the classifier for every pair of condition (instead of correlations between patterns of conditions) can then also be saved in an RDM. On a final note, different types of classifiers exist. Some simple linear classifiers include support-vector machine (SVM) and linear discriminant analysis (LDA). Non-linear classifiers could for example be artificial neural networks. In chapter 2 and 3 of this thesis, the cross-validated Mahalanobis distance classifier is used.

No matter how you approach the MVPA, a possible step afterwards is representational similarity analysis or RSA, where the MVPA results of one ROI can be compared to that of another ROI, or to a theoretical model, or to behavioral results, and so on. This is a very versatile method that allows all kind of data to be compared to each other (last row of Figure 13) (Kriegeskorte et al., 2008).

Representational similarity analysis (RSA)

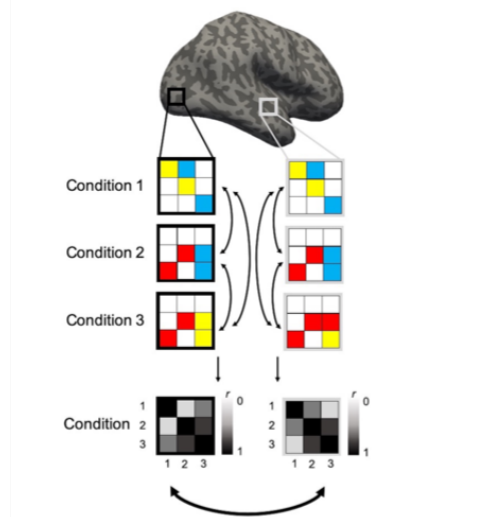


Figure 13. This figure visualizes two brain regions (one with a black outlined and one with a gray outlined square). The smaller squares inside are the voxels and the colors signify the neural activity (the beta values) of the voxels. On different rows, we find the activity of the voxels within the brain regions in response to different conditions. Patterns of different conditions can be correlated to each other and results are then saved in the black and white square on the bottom row, which is an RDM (correlational MVPA). You can do this for multiple ROIs, and then using representational similarity analysis, compare this result between ROIs. Adapted from Pillet and colleagues, 2019.

Functional connectivity analysis

In functional connectivity analysis the BOLD signal over time of voxels/ROIs is correlated to understand if the activity of these voxels/ROIs fluctuate together or not. The idea is that if two ROIs fluctuate together, they probable interact with one another and are thus functionally connected. For example, suppose that the amygdala and prefrontal cortex are connected during an emotion processing task. This would imply that the cognitive control center of the brain, the prefrontal cortex, is influencing the amygdala, which supports emotions processing. This method is used in chapter 1 of this thesis.

Functional connectivity analysis (FCA)

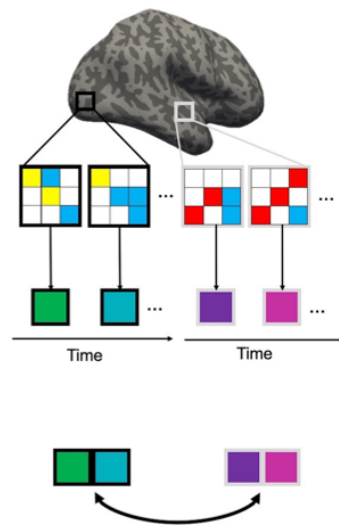


Figure 14. This figure visualizes two brain regions (one with a black outlined and one with a gray outlined square). The smaller squares inside are the voxels and the colors signify the neural activity (the beta values) of the voxels. We look at the neural activity, averaged across voxels, over time for both regions and can then correlate between regions. Adapted from Pillet and colleagues, 2019.

7T fMRI

Lately, a stronger magnet than 3T has become available for research: a 7T magnet. This is important as the strength of the magnet influences the spatial resolution by which the brain is scanned. The higher the resolution, the more spatial details we can see. This is not the only advantage, the quality of the signal also becomes better: the signal-to-noise ratio (SNR) increases (van der Zwaag et al., 2016). 7T fMRI is however somewhat challenging. It is more expensive, less stable, the space inside the scanner is smaller and thus it is more uncomfortable for the participant, and so on. Improvements and solutions are constantly being developed though (van der Zwaag et al., 2016).

7T fMRI has been proven to provide significant advances in the field of vision research by capitalizing upon the benefits that 7T fMRI provides (e.g., Yacoub et al., 2008; Olman & Yacoub, 2011; Nasr et al., 2016; van der Zwaag et al., 2016; Dumoulin et al., 2018; Schneider et al., 2019; Cai et al., 2021). McGugin and colleagues (2012) showed that using a high spatial resolution may completely change findings that were found using a lower spatial resolution. Even if an effect is not dependent on a high spatial resolution (see e.g., Margalit et al., 2020), the fact that 7T fMRI provides a higher SNR allows one to pick on smaller effects that could be picked up on before using data of a lower SNR from 3T fMRI. In fact, to pick up on certain neural findings, 7T fMRI may be a necessity unless one has the time, money and willing participants, to collect at least four times more data with a 3T fMRI (Cai et al., 2021). One can thus also choose, when using 7T fMRI, to collect less data per condition (when an effect

is large enough) and instead include more conditions in their experiment, which allows more conclusions to be made.

A recent study that showed how beneficial 7T fMRI can be specifically for investigating the functional organization of OTC, is the study by Boring and colleagues (2021). Boring and colleagues (2021) defined word and face areas using invasive intracranial recordings in multiple participants. They also performed 7T fMRI to locate the word and face areas. Results showed that with 7T fMRI, they could replicate the areas as found through the intracranial recordings. This demonstrates how 7T fMRI can capture the details of the functional organization of OTC, without the need for invasive brain imaging methods.

References

- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to “building” stimuli: Evidence and implications. *Neuron*, 21(2), 373–383.
[https://doi.org/10.1016/S0896-6273\(00\)80546-2](https://doi.org/10.1016/S0896-6273(00)80546-2)
- Almabruk, A. A., Paterson, K. B., McGowan, V., & Jordan, T. R. (2011). Evaluating effects of divided hemispheric processing on word recognition in foveal and extrafoveal displays: the evidence from Arabic. *PloS One*, 6(4), e18131. <https://doi.org/10.1371/journal.pone.0018131>
- Amaral, L., Bergström, F., & Almeida, J. (2021). Overlapping but distinct: distal connectivity dissociates hand and tool processing networks. *Cortex*, 140, 1–13.
<https://doi.org/10.1016/j.cortex.2021.03.011>
- Amedi, A., Stern, W. M., Camprodon, J. A., Bermpohl, F., Merabet, L., Rotman, S., Hemond, C., Meijer, P., & Pascual-Leone, A. (2007). Shape conveyed by visual-to-auditory sensory substitution activates the lateral occipital complex. *Nature Neuroscience*, 10(6), 687–689.
<https://doi.org/10.1038/nn1912>
- Arcaro, M. J., & Livingstone, M. S. (2017). A hierarchical, retinotopic proto-organization of the primate visual system at birth. *eLife*, 6, e26196. <https://doi.org/10.7554/eLife.26196>
- Arcaro, M. J., Schade, P. F., & Livingstone, M. S. (2019). Universal mechanisms and the development of the face network: what you see is what you get. *Annual Review of Vision Science*, 5, 341–372. <https://doi.org/10.1146/annurev-vision-091718-014917>
- Bach, P., Peelen, M. V., & Tipper, S. P. (2010). On the role of object information in action observation: an fMRI study. *Cerebral Cortex*, 20(12), 2798–2809. <https://doi.org/10.1093/cercor/bhq026>
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of*

- the National Academy of Sciences of the United States of America*, 104(21), 9087–9092.
<https://doi.org/10.1073/pnas.0703300104>
- Bao, P., She, L., McGill, M., & Tsao, D. Y. (2020). A map of object space in primate inferotemporal cortex. *Nature*, 583(7814), 103–108. <https://doi.org/10.1038/s41586-020-2350-5>
- Barton, J. J. S., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, 58(1), 71–78.
<https://doi.org/10.1212/WNL.58.1.71>
- Beauchamp M. S. (2005). See me, hear me, touch me: multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, 15(2), 145–153.
<https://doi.org/10.1016/j.conb.2005.03.011>
- Bedny M. (2017). Evidence from blindness for a cognitively pluripotent cortex. *Trends in Cognitive Sciences*, 21(9), 637–648. <https://doi.org/10.1016/j.tics.2017.06.003>
- Behrmann, M., & Plaut, D. C. (2020). Hemispheric organization for visual object recognition: a theoretical account and empirical evidence. *Perception*, 49(4), 373–404.
<https://doi.org/10.1177/0301006619899049>
- Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2007). Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cerebral Cortex*, 17(7), 1604–1611.
<https://doi.org/10.1093/cercor/bhl071>
- Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., Bulthé, J., Sunaert, S., Wouters, J., & Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science*, 342(6163), 1251–1254.
<https://doi.org/10.1126/science.1244333>
- Boring, M. J., Silson, E. H., Ward, M. J., Richardson, R. M., Fiez, J. A., Baker, C. I., & Ghuman, A. S. (2021). Multiple adjoining word- and face-selective regions in ventral temporal cortex exhibit distinct dynamics. *The Journal of Neuroscience*, 41(29), 6314–6327.
<https://doi.org/10.1523/JNEUROSCI.3234-20.2021>
- 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.
<https://doi.org/10.1523/JNEUROSCI.2314-15.2016>
- Bracci, S., Cavina-Pratesi, C., Connolly, J. D., & Ietswaart, M. (2016). Representational content of occipitotemporal and parietal tool areas. *Neuropsychologia*, 84, 81–88.
<https://doi.org/10.1016/j.neuropsychologia.2015.09.001>

- Bracci, S., & Peelen, M. V. (2013). Body and object effectors: the organization of object representations in high-level visual cortex reflects body-object interactions. *The Journal of Neuroscience*, 33(46), 18247–18258. <https://doi.org/10.1523/JNEUROSCI.1322-13.2013>
- 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. <https://doi.org/10.1152/jn.00619.2011>
- 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. <https://doi.org/10.1152/jn.00215.2010>
- Brants, M., Bulthé, J., Daniels, N., Wagemans, J., & Beeck, H.O. (2016). How learning might strengthen existing visual object representations in human object-selective cortex. *NeuroImage*, 127, 74-85.
- Budisavljevic, S., Dell'Acqua, F., & Castiello, U. (2018). Cross-talk connections underlying dorsal and ventral stream integration during hand actions. *Cortex*, 103, 224–239. <https://doi.org/10.1016/j.cortex.2018.02.016>
- Caffarra, S., Karipidis, I. I., Yablonski, M., & Yeatman, J. D. (2021). Anatomy and physiology of word-selective visual cortex: from visual features to lexical processing. *Brain Structure and Function*, 226(9), 3051–3065. <https://doi.org/10.1007/s00429-021-02384-8>
- Cai, Y., Hofstetter, S., van der Zwaag, W., Zuiderbaan, W., & Dumoulin, S. O. (2021). Individualized cognitive neuroscience needs 7T: Comparing numerosity maps at 3T and 7T MRI. *NeuroImage*, 237(118184), 118184. <https://doi.org/10.1016/j.neuroimage.2021.118184>
- 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. <https://doi.org/10.1038/13217>
- Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *NeuroImage*, 12(4), 478–484. <https://doi.org/10.1006/nimg.2000.0635>
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 362(1485), 1585–1599. <https://doi.org/10.1098/rstb.2007.2054>
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, 123, 291–307. <https://doi.org/10.1093/brain/123.2.291>

- Cohen, L., Stéphane, L., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, 125(5), 1054-1069. <https://doi.org/10.1093/brain/awf094>
- Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *NeuroImage*, 22(1), 466–476. <https://doi.org/10.1016/j.neuroimage.2003.12.049>
- Coggan, D. D., Liu, W., Baker, D. H., & Andrews, T. J. (2016). Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. *NeuroImage*, 135, 107–114. <https://doi.org/10.1016/j.neuroimage.2016.04.060>
- Dean P. (1976). Effects of inferotemporal lesions on the behavior of monkeys. *Psychological Bulletin*, 83(1), 41–71.
- Deen, B., Richardson, H., Dilks, D. D., Takahashi, A., Keil, B., Wald, L. L., Kanwisher, N., & Saxe, R. (2017). Organization of high-level visual cortex in human infants. *Nature Communications*, 8, 13995. <https://doi.org/10.1038/ncomms13995>
- Dehaene, S., Le Clec'h, G., Poline, J.-B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, 13(3), 321–325. <https://doi.org/10.1097/00001756-200203040-00015>
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, 9(7), 335–341. <https://doi.org/10.1016/j.tics.2005.05.004>
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384–398. <https://doi.org/10.1016/j.neuron.2007.10.004>
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, 330(6009), 1359–1364. <https://doi.org/10.1126/science.1194140>
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15(6), 254–262. <https://doi.org/10.1016/j.tics.2011.04.003>
- Dehaene-Lambertz, G., Monzalvo, K., & Dehaene, S. (2018). The emergence of the visual word form: longitudinal evolution of category-specific ventral visual areas during reading acquisition. *PLoS Biology*, 16(3), e2004103. <https://doi.org/10.1371/journal.pbio.2004103>
- Downing, P., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area specialized for visual processing of the human body. *Science*, 293, 2470–2473. <https://doi.org/10.1126/science.1063414>

- Dumoulin, S. O., Fracasso, A., van der Zwaag, W., Siero, J. C. W., & Petridou, N. (2018). Ultra-high field MRI: Advancing systems neuroscience towards mesoscopic human brain function. *NeuroImage*, 168, 345–357. <https://doi.org/10.1016/j.neuroimage.2017.01.028>
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392, 598–601. <https://doi.org/10.1038/33402>
- Fairhall, S. L., Porter, K. B., Bellucci, C., Mazzetti, M., Cipolli, C., & Gobbini, M. I. (2017). Plastic reorganization of neural systems for perception of others in the congenitally blind. *NeuroImage*, 158, 126–135. <https://doi.org/10.1016/j.neuroimage.2017.06.057>
- Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision*. The MIT Press.
- Felleman, D.J. and Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47. <http://dx.doi.org/10.1093/cercor/1.1.1>
- Freud, E., Culham, J. C., Plaut, D. C., & Behrmann, M. (2017). The large-scale organization of shape processing in the ventral and dorsal pathways. *eLife*, 6. <https://doi.org/10.7554/elife.27576>
- Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., ... Cohen, L. (2006). Direct intracranial, fmri, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron*, 50(2), 191–204. <https://doi.org/10.1016/j.neuron.2006.03.031>
- Garcea, F. E., & Mahon, B. Z. (2014). Parcellation of left parietal tool representations by functional connectivity. *Neuropsychologia*, 60, 131–143. <https://doi.org/10.1016/j.neuropsychologia.2014.05.018>
- Gerrits, R., Van der Haegen, L., Brysbaert, M., & Vingerhoets, G. (2019). Laterality for recognizing written words and faces in the fusiform gyrus covaries with language dominance. *Cortex*, 117, 196–204. <https://doi.org/10.1016/j.cortex.2019.03.010>
- Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature reviews Neuroscience*, 14(5), 350–363. <https://doi.org/10.1038/nrn3476>
- Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., & Fagot, J. (2012). Orthographic processing in baboons (*Papio papio*). *Science*, 336(6078), 245–248. <https://doi.org/10.1126/science.1218152>
- Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to “mirror” and “canonical” neurones in the human brain: an fMRI study. *NeuroImage*, 18(4), 928–937. [https://doi.org/10.1016/s1053-8119\(03\)00042-9](https://doi.org/10.1016/s1053-8119(03)00042-9)
- Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, 3(8), 837– 843. <https://doi.org/10.1038/77754>

- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, 27(1), 649–677. <https://doi.org/10.1146/annurev.neuro.27.070203.144220>
- Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition: as soon as you know it is there, you know what it is. *Psychological Science*, 16(2), 152–160. <https://doi.org/10.1111/j.0956-7976.2005.00796.x>
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, 15(8), 536–548. <https://doi.org/10.1038/nrn3747>
- Grill-Spector, K., Weiner, K. S., Kay, K., & Gomez, J. (2017). The functional neuroanatomy of human face perception. *Annual Review of Vision Science*, 3(1), 167–196. <https://doi.org/10.1146/annurev-vision-102016-061214>
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of face and objects in ventral temporal cortex. *Science*, 293, 2425–2431. <https://doi.org/10.1126/science.1063736>
- 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. [https://doi.org/10.1016/S0896-6273\(02\)00662-1](https://doi.org/10.1016/S0896-6273(02)00662-1)
- Henly (n.d.). <https://openbooks.lib.msu.edu/app/uploads/sites/6/2021/03/VisualStreams.png>
- Hervais-Adelman, A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., Eisner, F., & Huettig, F. (2019). Learning to read recycles visual cortical networks without destruction. *Science Advances*, 5(9), eaax0262. <https://doi.org/10.1126/sciadv.aax0262>
- Huey, E. B. (1900). On the Psychology and Physiology of Reading. I. *The American Journal of Psychology*, 11(3), 283–302. <https://doi.org/10.2307/1412745>
- Janzen, G., & van Turenhout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature Neuroscience*, 7(6), 673–677. <https://doi.org/10.1038/nn1257>
- Kaiser, D., Azzalini, D. C., & Peelen, M. V. (2016). Shape-independent object category responses revealed by MEG and fMRI decoding. *Journal of Neurophysiology*, 115(4), 2246–2250. <https://doi.org/10.1152/jn.01074.2015>
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11), 4302–4311. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>
- 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 of the United States of America*, 107(25), 11163–11170. <https://doi.org/10.1073/pnas.1005062107>

- Kim, J. S., Kanjlia, S., Merabet, L. B., & Bedny, M. (2017). Development of the visual word form area requires visual experience: evidence from blind braille readers. *The Journal of Neuroscience*, 37(47), 11495–11504. <https://doi.org/10.1523/JNEUROSCI.0997-17.2017>
- Knights, E., Mansfield, C., Tonin, D., Saada, J., Smith, F. W., & Rossit, S. (2021). Hand-selective visual regions represent how to grasp 3D tools: brain decoding during real actions. *The Journal of Neuroscience*, 41(24), 5263–5273. <https://doi.org/10.1523/JNEUROSCI.0083-21.2021>
- Konen, C. S., Behrmann, M., Nishimura, M., & Kastner, S. (2011). The functional neuroanatomy of object agnosia: a case study. *Neuron*, 71(1), 49–60. <https://doi.org/10.1016/j.neuron.2011.05.030>
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, 74(6), 1114–1124. <https://doi.org/10.1016/j.neuron.2012.04.036>
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2, 4. <https://doi.org/10.3389/neuro.06.004.2008>
- Kubota, E. C., Joo, S. J., Huber, E., & Yeatman, J. D. (2019). Word selectivity in high-level visual cortex and reading skill. *Developmental Cognitive Neuroscience*, 36, 100593. <https://doi.org/10.1016/j.dcn.2018.09.003>
- Le, R., Witthoft, N., Ben-Shachar, M., & Wandell, B. (2017). The field of view available to the ventral occipito-temporal reading circuitry. *Journal of Vision*, 17(4), 1–19. <https://doi.org/10.1167/17.4.6>
- Legge, G. E., Mansfield, J. S., & Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research*, 41(6), 725–743. [https://doi.org/10.1016/s0042-6989\(00\)00295-9](https://doi.org/10.1016/s0042-6989(00)00295-9)
- Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences of the United States of America*, 115(42), E9981–E9990. <https://doi.org/10.1073/pnas.1803003115>
- Lerma-Usabiaga, G., Le, R., Gafni, C., Ben-Shachar, M., & Wandell, B. (2021). Interpreting sensory and cognitive signals in the cortical reading network. *bioRxiv*. <https://doi.org/10.1101/2021.09.14.460238>
- Leshinskaya, A., Bajaj, M., & Thompson-Schill, S. L. (2020). Tool-selective lateral temporal cortex responds to objects with causal effects. *Journal of Vision*, 20(11), 1417. <https://doi.org/10.1167/jov.20.11.1417>

- Livingstone, M. S., Vincent, J. L., Arcaro, M. J., Srihasam, K., Schade, P. F., & Savage, T. (2017). Development of the macaque face-patch system. *Nature Communications*, 8, 14897. <https://doi.org/10.1038/ncomms14897>
- Logothetis, N. K. (2002). The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 357(1424), 1003–1037. <https://doi.org/10.1098/rstb.2002.1114>
- Magri, C., Konkle, T., & Caramazza, A. (2016). Visual object responses of the ventral stream reflect both size and motor-relevance. *Journal of Vision*, 16(12), 505. <https://doi.org/10.1167/16.12.505>
- Mahon, B. Z., Anzellotti, S., Schwarzbach, J., Zampini, M., & Caramazza, A. (2009). Category-specific organization in the human brain does not require visual experience. *Neuron*, 63(3), 397–405. <https://doi.org/10.1016/j.neuron.2009.07.012>
- Mahon, B. Z., & Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. *Annual Review of Psychology*, 60(1), 27–51. <https://doi.org/10.1146/annurev.psych.60.110707.163532>
- Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain? *Trends in Cognitive Sciences*, 15(3), 97–103. <https://doi.org/10.1016/j.tics.2011.01.004>
- Mahon, B. Z., Kumar, N., & Almeida, J. (2013). Spatial frequency tuning reveals interactions between the dorsal and ventral visual systems. *Journal of Cognitive Neuroscience*, 25(6), 862–871. https://doi.org/10.1162/jocn_a_00370
- Mahon, B. Z., Milleville, S. C., Negri, G. A. L., Rumiat, R. I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, 55(3), 507–520. <https://doi.org/10.1016/j.neuron.2007.07.011>
- Maimon-Mor, R. O., & Makin, T. R. (2020). Is an artificial limb embodied as a hand? Brain decoding in prosthetic limb users. *PLoS Biology*, 18(6), e3000729. <https://doi.org/10.1371/journal.pbio.3000729>
- Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences*, 6(4), 176–184. [https://doi.org/10.1016/s1364-6613\(02\)01870-3](https://doi.org/10.1016/s1364-6613(02)01870-3)
- Margalit, E., Jamison, K. W., Weiner, K. S., Vizioli, L., Zhang, R.-Y., Kay, K. N., & Grill-Spector, K. (2020). Ultra-high-resolution fMRI of human ventral temporal cortex reveals differential representation of categories and domains. *The Journal of Neuroscience*, 40(15), 3008–3024. <https://doi.org/10.1523/JNEUROSCI.2106-19.2020>

- Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco, CA: W.H. Freeman.
- Mattioni, S., Rezk, M., Battal, C., Bottini, R., Cuculiza Mendoza, K. E., Oosterhof, N. N., & Collignon, O. (2020). Categorical representation from sound and sight in the ventral occipito-temporal cortex of sighted and blind. *eLife*, 9, e50732. <https://doi.org/10.7554/eLife.50732>
- McGugin, R. W., Gatenby, J. C., Gore, J. C., & Gauthier, I. (2012). High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. *Proceedings of the National Academy of Sciences of the United States of America*, 109(42), 17063–17068. <https://doi.org/10.1073/pnas.1116333109>
- McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: experience is less important than previously believed. *Cognitive Neuropsychology*, 29(1-2), 174–212. <https://doi.org/10.1080/02643294.2012.660138>
- Mishkin, M., & Gorgays, D. G. (1952). Word recognition as a function of retinal locus. *Journal of Experimental Psychology*, 43(1), 43–48. <https://doi.org/10.1037/h0061361>
- Moore, M. W., Durisko, C., Perfetti, C. A., & Fiez, J. A. (2014). Learning to read an alphabet of human faces produces left-lateralized training effects in the fusiform gyrus. *Journal of Cognitive Neuroscience*, 26(4), 896–913. https://doi.org/10.1162/jocn_a_00506
- Moutoussis, K., & Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14), 9527–9532. <https://doi.org/10.1073/pnas.142305699>
- Nasr, S., Polimeni, J. R., & Tootell, R. B. H. (2016). Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. *The Journal of Neuroscience*, 36(6), 1841–1857. <https://doi.org/10.1523/JNEUROSCI.3518-15.2016>
- Nieto, 2015. https://commons.wikimedia.org/wiki/File:Human_visual_pathway.svg
- Nordt, M., Gomez, J., Natu, V. S., Rezai, A. A., Finzi, D., Kular, H., & Grill-Spector, K. (2021). Cortical recycling in high-level visual cortex during childhood development. *Nature Human Behaviour*, 5(12), 1686–1697. <https://doi.org/10.1038/s41562-021-01141-5>
- O’Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience*, 12(6), 1013–1023. <https://doi.org/10.1162/08989290051137549>
- Olman, C. A., & Yacoub, E. (2011). High-field fMRI for human applications: an overview of spatial resolution and signal specificity. *The Open Neuroimaging Journal*, 5(Suppl 1), 74–89. <https://doi.org/10.2174/18744440001105010074>

- Op De Beeck, H. P., 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. <https://doi.org/10.1523/JNEUROSCI.2481-06.2006>
- Op de Beeck, H. P., & Baker, C. I. (2010). The neural basis of visual object learning. *Trends in Cognitive Sciences*, 14(1), 22–30. <https://doi.org/10.1016/j.tics.2009.11.002>
- Op de Beeck, H. P., Pillet, I., & Ritchie, J. B. (2019). Factors determining where category-selective areas emerge in visual cortex. *Trends in Cognitive Sciences*, 23(9), 784–797. <https://doi.org/10.1016/j.tics.2019.06.006>
- Orlov, T., Makin, T. R., & Zohary, E. (2010). Topographic representation of the human body in the occipitotemporal cortex. *Neuron*, 68(3), 586–600. <https://doi.org/10.1016/j.neuron.2010.09.032>
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, 93(1), 603–608. <https://doi.org/10.1152/jn.00513.2004>
- Peelen, M. V., & Caramazza, A. (2010). What body parts reveal about the organization of the brain. *Neuron*, 68(3), 331–333. <https://doi.org/10.1016/j.neuron.2010.10.021>
- 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. https://doi.org/10.1162/jocn_a_00411
- Peeters, R. R., Rizzolatti, G., & Orban, G. A. (2013). Functional properties of the left parietal tool use region. *NeuroImage*, 78, 83–93. <https://doi.org/10.1016/j.neuroimage.2013.04.023>
- Perini, F., Caramazza, A., & Peelen, M. V. (2014). Left occipitotemporal cortex contributes to the discrimination of tool-associated hand actions: fMRI and TMS evidence. *Frontiers in Human Neuroscience*, 8, 591. <https://doi.org/10.3389/fnhum.2014.00591>
- Pietrini, P., Furey, M. L., Ricciardi, E., Gobbin, M. I., Wu, W.-H. C., Cohen, L., Guazzelli, M., & Haxby, J. V. (2004). Beyond sensory images: object-based representation in the human ventral pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 101(15), 5658–5663. <https://doi.org/10.1073/pnas.0400707101>
- Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *The Journal of Neuroscience*, 16(16), 5205–5215. <https://doi.org/10.1523/jneurosci.16-16-05205.1996>
- Proklova, D., Kaiser, D., & Peelen, M. V. (2016). Disentangling representations of object shape and object category in human visual cortex: the animate–inanimate distinction. *Journal of Cognitive Neuroscience*, 28(5), 680–692. https://doi.org/10.1162/jocn_a_00924

- Rajalingham, R., Kar, K., Sanghavi, S., Dehaene, S., & DiCarlo, J. J. (2020) The inferior temporal cortex is a potential cortical precursor of orthographic processing in untrained monkeys. *Nature Communications*, 11, 3886. <https://doi.org/10.1038/s41467-020-17714-3>
- Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012). Position sensitivity in the visual word form area. *Proceedings of the National Academy of Sciences of the United States of America*, 109(24), E1568–E1577. <https://doi.org/10.1073/pnas.1121304109>
- Rayner K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Current Biology*, 21(5), 363–368. <https://doi.org/10.1016/j.cub.2011.01.040>
- Rice, G. E., Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. *The Journal of Neuroscience*, 34(26), 8837–8844. <https://doi.org/10.1523/JNEUROSCI.5265-13.2014>
- Richlan, F., Kronbichler, M., & Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *NeuroImage*, 56(3), 1735–1742. <https://doi.org/10.1016/j.neuroimage.2011.02.040>
- Sakuraba, S., Sakai, S., Yamanaka, M., Yokosawa, K., & Hirayama, K. (2012). Does the human dorsal stream really process a category for tools? *The Journal of Neuroscience*, 32(11), 3949–3953. <https://doi.org/10.1523/JNEUROSCI.3973-11.2012>
- Sathian, K., & Lacey, S. (2007). Journeying beyond classical somatosensory cortex. *Revue Canadienne de Psychologie Experimentale [Canadian Journal of Experimental Psychology]*, 61(3), 254–264. <https://doi.org/10.1037/cjep2007026>
- Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D., & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, 19(9), 1250–1255. <https://doi.org/10.1038/nn.4354>
- Schneider, M., Kemper, V. G., Emmerling, T. C., De Martino, F., & Goebel, R. (2019). Columnar clusters in the human motion complex reflect consciously perceived motion axis. *Proceedings of the National Academy of Sciences of the United States of America*, 116(11), 5096–5101. <https://doi.org/10.1073/pnas.1814504116>

- Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., & Rossion, B. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cerebral Cortex*, 16(4), 574–586.
<https://doi.org/10.1093/cercor/bhj005>
- Schone, H. R., Maimon-Mor, R. O., Baker, C. I., & Makin, T. R. (2021). Expert tool users show increased differentiation between visual representations of hands and tools. *The Journal of Neuroscience*, 41(13), 2980–2989. <https://doi.org/10.1523/JNEUROSCI.2489-20.2020>
- Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences*, 7(11), 489–493.
<https://doi.org/10.1016/j.tics.2003.09.010>
- Serrien, D. J., Ivry, R. B., & Swinnen, S. P. (2006). Dynamics of hemispheric specialization and integration in the context of motor control. *Nature Reviews. Neuroscience*, 7(2), 160–166.
<https://doi.org/10.1038/nrn1849>
- Siéoroff, E., & Haehnel-Benoliel, N. (2015). Environmental script affects lateral asymmetry of word recognition: a study of French-Hebrew bilinguals tested in Israel and in France. *Laterality*, 20(4), 389–417. <https://doi.org/10.1080/1357650X.2014.988220>
- Srihasam, K., Mandeville, J. B., Morocz, I. A., Sullivan, K. J., & Livingstone, M. S. (2012). Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron*, 73(3), 608–619. <https://doi.org/10.1016/j.neuron.2011.12.022>
- Striem-Amit, E., Cohen, L., Dehaene, S., & Amedi, A. (2012). Reading with sounds: sensory substitution selectively activates the visual word form area in the blind. *Neuron*, 76(3), 640–652. <https://doi.org/10.1016/j.neuron.2012.08.026>
- Striem-Amit, E., Vannuscorps, G., & Caramazza, A. (2017). Sensorimotor-independent development of hands and tools selectivity in the visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 114(18), 4787–4792.
<https://doi.org/10.1073/pnas.1620289114>
- Tarhan, L. Y., Watson, C. E., & Buxbaum, L. J. (2015). Shared and distinct neuroanatomic regions critical for tool-related action production and recognition: evidence from 131 left-hemisphere stroke patients. *Journal of Cognitive Neuroscience*, 27(12), 2491–2511.
https://doi.org/10.1162/jocn_a_00876
- Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522. <https://doi.org/10.1038/381520a0>

- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*, 670–674.
<https://doi.org/10.1126/science.1119983>
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549-586). Cambridge: MIT Press.
- van den Heiligenberg, F. M. Z., Orlov, T., Macdonald, S. N., Duff, E. P., Henderson Slater, D., Beckmann, C. F., Johansen-Berg, H., Culham, J. C., & Makin, T. R. (2018). Artificial limb representation in amputees. *Brain*, *141*(5), 1422–1433. <https://doi.org/10.1093/brain/awy054>
- van den Heiligenberg, F., Macdonald, S., Duff, E., Henderson Slater, D., Johansen-Berg, H., Culham, J., & Makin, T. (2015). Activity in hand- and tool-selective regions for prosthetic limbs in amputees is associated with prosthesis usage in everyday life. *Journal of Vision*, *15*(12), 983. <https://doi.org/10.1167/15.12.983>
- van den Hurk, J., Van Baelen, M., & Op de Beeck, H. P. (2017). Development of visual category selectivity in ventral visual cortex does not require visual experience. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(22), E4501–E4510. <https://doi.org/10.1073/pnas.1612862114>
- Van der Haegen, L., Cai, Q., & Brysbaert, M. (2012). Colateralization of Broca's area and the visual word form area in left-handers: fMRI evidence. *Brain and Language*, *122*(3), 171–178. <https://doi.org/10.1016/j.bandl.2011.11.004>
- van der Zwaag, W., Schäfer, A., Marques, J. P., Turner, R., & Trampel, R. (2016). Recent applications of UHF-MRI in the study of human brain function and structure: a review: UHF MRI: Applications to Human Brain Function and Structure. *NMR in Biomedicine*, *29*(9), 1274–1288. <https://doi.org/10.1002/nbm.3275>
- Van Essen, D. C. (2003). Organization of visual areas in macaque and human cerebral cortex. In L. Chalupa & J. Werner (Eds.), *The Visual Neurosciences* (pp. 507-521). Cambridge, MA: MIT Press.
- Valyear, K. F., Cavina-Pratesi, C., Stiglick, A. J., & Culham, J. C. (2007). Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *NeuroImage*, *36 Suppl 2*, T94–T108. <https://doi.org/10.1016/j.neuroimage.2007.03.031>
- Valyear, K. F., & Culham, J. C. (2010). Observing learned object-specific functional grasps preferentially activates the ventral stream. *Journal of Cognitive Neuroscience*, *22*(5), 970–984. <https://doi.org/10.1162/jocn.2009.21256>

- Vogel, A. C., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2012). The putative visual word form area is functionally connected to the dorsal attention network. *Cerebral Cortex*, 22(3), 537–549. <https://doi.org/10.1093/cercor/bhr100>
- Wandell, B. A., Rauschecker, A. M., & Yeatman, J. D. (2012). Learning to see words. *Annual Review of Psychology*, 63, 31–53. <https://doi.org/10.1146/annurev-psych-120710-100434>
- Watson, D. M., Hartley, T., & Andrews, T. J. (2014). Patterns of response to visual scenes are linked to the low-level properties of the image. *NeuroImage*, 99, 402–410. <https://doi.org/10.1016/j.neuroimage.2014.05.045>
- Weiner, K. S., & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage*, 52(4), 1559–1573. <https://doi.org/10.1016/j.neuroimage.2010.04.262>
- Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., ... Grill-Spector, K. (2014). The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage*, 84, 453–465. <https://doi.org/10.1016/j.neuroimage.2013.08.068>
- Weisberg, J., van Turennout, M., & Martin, A. (2007) A neural system for learning about object function. *Cerebral Cortex*, 17, 513–521. <http://dx.doi.org/10.1093/cercor/bhj176>
- White, A. L., Palmer, J., Boynton, G. M., & Yeatman, J. D. (2019). Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proceedings of the National Academy of Sciences*, 116(20), 10087–10096. <https://doi.org/10.1073/pnas.1822137116>
- White, A. L., Boynton, G. M., & Yeatman, J. D. (2019)b. The link between reading ability and visual spatial attention across development. *Cortex*, 121, 44–59. <https://doi.org/10.1016/j.cortex.2019.08.011>
- Willems, R. M., Peelen, M. V., & Hagoort, P. (2010). Cerebral lateralization of face-selective and body-selective visual areas depends on handedness. *Cerebral Cortex*, 20(7), 1719–1725. <https://doi.org/10.1093/cercor/bhp234>
- Wong, A. C., Palmeri, T. J., Rogers, B. P., Gore, J. C., & Gauthier, I. (2009). Beyond shape: how you learn about objects affects how they are represented in visual cortex. *PloS One*, 4(12), e8405. <https://doi.org/10.1371/journal.pone.0008405>
- Yacoub, E., Harel, N., & Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(30), 10607–10612. <https://doi.org/10.1073/pnas.0804110105>

- Yargholi, E., & Op de Beeck, H. (2023). Category trumps shape as an organizational principle of object space in the human occipitotemporal cortex. *Journal of Neuroscience*, 43(16), 2960-2972. <https://doi.org/10.1523/JNEUROSCI.2179-22.2023>
- Yeatman, J. D., White, A. L., Yeatman, J. D., & Mall, G. (2021). Reading: the confluence of vision and language. *Annual Review of Vision Science*, 7, 487–517. <https://doi.org/10.1146/annurev-vision-093019-113509>
- Zabaleta (2019). https://commons.wikimedia.org/wiki/File:Drawing_of_a_neuron.svg
- Zeki S. (2016). Multiple asynchronous stimulus- and task-dependent hierarchies (STDH) within the visual brain's parallel processing systems. *The European Journal of Neuroscience*, 44(8), 2515–2527. <https://doi.org/10.1111/ejn.13270>
- [Lobes of the cerebral cortex] (n.d.). https://mhanational.org/sites/default/files/Picture1_0.png
- [Gray and white matter] (n.d.). <https://assets.technologynetworks.com/production/dynamic/images/content/322973/gray-matter-vs-white-matter-322973-1280x720.webp?cb=12181651>

Chapter 1: A comparison of functional networks derived from representational similarity, functional connectivity, and univariate analyses

This chapter as provided in this thesis has been published. Reference: Pillet, I., Op de Beeck, H., & Lee Masson, H. (2020). A comparison of functional networks derived from representational similarity, functional connectivity, and univariate analyses. *Frontiers in Neuroscience*, 13, 1348. <https://doi.org/10.3389/fnins.2019.01348>

Abstract

The invention of representational similarity analysis [RSA, following multi-voxel pattern analysis (MVPA)] has allowed cognitive neuroscientists to identify the representational structure of multiple brain regions, moving beyond functional localization. By comparing these structures, cognitive neuroscientists can characterize how brain areas form functional networks. Univariate analysis (UNIVAR) and functional connectivity analysis (FCA) are two other popular methods to identify functional networks. Despite their popularity, few studies have examined the relationship between networks from RSA with those from UNIVAR and FCA. Thus, the aim of the current study is to examine the similarities between neural networks derived from RSA with those from UNIVAR and FCA to explore how these methods relate to each other. We analyzed the data of a previously published study with the three methods and compared the results by performing (partial) correlation and multiple regression analysis. Our findings reveal that neural networks resulting from RSA, UNIVAR, and FCA methods are highly similar to each other even after ruling out the effect of anatomical proximity between the network nodes. Nevertheless, the neural network from each method shows unique organization that cannot be explained by any of the other methods. Thus, we conclude that the RSA, UNIVAR and FCA methods provide similar but not identical information on how brain regions are organized in functional networks.

Introduction

Multi-voxel pattern analysis (MVPA) has recently become one of the most frequently used techniques for analyzing fMRI data. It considers the spatial pattern of neural activation across multiple voxels and examines whether these patterns contain task-related information (Haxby et al., 2001; Haynes and Rees, 2006; Mur et al., 2009; Kriegeskorte, 2011; Haxby, 2012; Coutanche, 2013; Haynes, 2015). It is referred to as *multivariate* or *multi-voxel* because it analyzes a set of voxels together (the pattern of activation of this set) instead of modeling activity of a single voxel (as is done in univariate analysis) (Norman et al., 2006; Mur et al., 2009; Kriegeskorte, 2011; Yang et al., 2012). In addition, patterns of activation can be used to investigate the similarities between such patterns of different conditions, for example the similarity of the activation pattern when viewing a face vs. when viewing a scene, or between such activation patterns of different brain regions in a certain condition (Mur et al., 2009; Haxby, 2012). This approach is referred to as representational similarity analysis (RSA) (Kriegeskorte et al., 2008). In first-order RSA, a representational dissimilarity matrix (RDM) is set up to understand the dissimilarity between patterns of activation of different stimuli in a certain brain region (Kriegeskorte et al., 2008; Yang et al., 2012). In second-order RSA, RDMs are compared between brain regions (Kriegeskorte et al., 2008; Yang et al., 2012). This method has been referred to as representational connectivity as it allows to identify the representational relationship among brain regions (Kriegeskorte et al., 2008). Connectivity related to multivariate information has since then been given a more specific meaning to refer to analyses of the temporal dynamics of the information contained in multi-voxel patterns, also sometimes referred to as multivariate or informational connectivity (Coutanche and Thompson-Schill, 2014; Anzellotti and Coutanche, 2018). For this reason, we opted for the more general RSA term instead of using the term representational connectivity.

Univariate analysis (UNIVAR) and functional connectivity analysis (FCA) are two other frequently used techniques for analyzing fMRI data. UNIVAR assesses neural activation of an individual voxel or a mean activation across voxels of a brain region. For this reason, it is often used to localize brain regions engaged in processing a particular type of stimuli (e.g., face vs. object) and thereby draw conclusions about the regions that are involved in cognitive processes important for the stimuli or task at hand (Haynes and Rees, 2006; Logothetis, 2008; Mur et al., 2009; Coutanche, 2013; Haynes, 2015). It is referred to as univariate because a general linear model (GLM) is applied voxel- wise to relate the experimental design to the neural activity of each voxel's time-course in the brain (Raizada and Kriegeskorte, 2010). FCA (for a review, see Friston, 2011) characterizes the communication between brain regions during rest or a task (Friston, 1994), measuring the strength of the relation between BOLD time-series signals of brain regions (Yang et al., 2012; Geerligs et al., 2016). When FCA is applied to a resting-state fMRI dataset, it reveals the intrinsic network of the brain based on low-frequency

BOLD fluctuations of brain regions (Biswal et al., 1995; Cordes et al., 2001; Fox and Raichle, 2007). This intrinsic network can also be extracted from a task-based fMRI dataset by removing the task-induced signal from the data (Fair et al., 2007). Therefore, this method is often referred to as *intrinsic functional connectivity*.

Notably, there are various conceptual similarities and differences between RSA, UNIVAR, and FCA. UNIVAR and FCA methods are similar in that they average across the BOLD signal of all the voxels in a brain region, unlike MVPA (see Figure 1). The organization of networks from co-activation has also proven to be similar to those from resting-state connectivity (Crossley et al., 2013). Analogously, Anzellotti and Coutanche referred to this type of FCA as univariate FCA (Anzellotti and Coutanche, 2018). Second-order RSA and FCA are similar in that they are both based on a measure of the similarity between brain regions. When using correlations, correlating the averaged BOLD time-series signals between the regions of interest (ROIs) in FCA is methodologically similar to correlating RDMs of those ROIs in second-order RSA (Xue et al., 2013). UNIVAR and RSA, or at least MVPA, have been frequently compared when describing functional properties of one region of the brain (e.g., see Jimura and Poldrack, 2012; Coutanche, 2013; Davis et al., 2014; Gilron et al., 2017). A significant finding from these studies was that changes (across different stimuli) in the activation patterns could be detected even when conditions were not different in the average univariate activation in a region (Mur et al., 2009). For example, different speech sounds showed different activation patterns in the right auditory cortex, but the average activation of this region across those speech sounds did not differ (Raizada et al., 2010). These studies have provided valuable insights into the conceptual and empirical relationships between UNIVAR and MVPA. Similarly, studies have used both RSA and FCA, some drawing the same conclusion from the results of RSA and FCA (e.g., Zeharia et al., 2015), or not (Boets et al., 2013; Bulthé et al., 2018).

A study that directly and simultaneously compares networks resulting from RSA with those from UNIVAR and FCA is missing from today's literature. Thus, the current study explores how the networks from RSA complement the networks from UNIVAR and FCA when investigating the functional architecture of the brain. Although direct and simultaneous comparisons between brain networks based on UNIVAR, RSA and FCA have not been performed (to our knowledge), we expect at least some convergence. For example, we hypothesize that brain regions with similar representational similarity would tend to be functionally connected, without excluding the possibility of uniqueness in the networks resulting from the two methods. Specifically, given the evidence of the topographic arrangement of the basic sensory cortical areas, such as the visual and sensorimotor cortex (see Kaas, 1997, for a detailed review) and given that these low-level visual or sensorimotor areas respectively are similar in what they do even at rest (Cordes et al., 2000), we predict that the way in which brain

networks composed of visual or sensorimotor areas are constructed, would be highly similar in all three methods. More precisely, we predict that the visual network organization and the sensorimotor network organization respectively, as obtained by UNIVAR, RSA, and FCA, will be highly similar. In sum, the goal of this study is to explore how the networks derived from RSA compare to those from UNIVAR and FCA. To answer this question, we applied second-order RSA to a previously reported fMRI study (Lee Masson et al., 2018) and compared the resulting networks with the results obtained from UNIVAR and intrinsic FCA. In particular, we conducted (partial) correlation and multiple regression analysis (comparing them to signal-to-noise ratio measurements), controlling for the confounding influence of anatomical proximity between brain regions of interest on RSA, UNIVAR, and intrinsic FCA results. In addition, we explored our results visually by implementing multi-dimensional scaling (MDS) and Procrustes transformation methods.

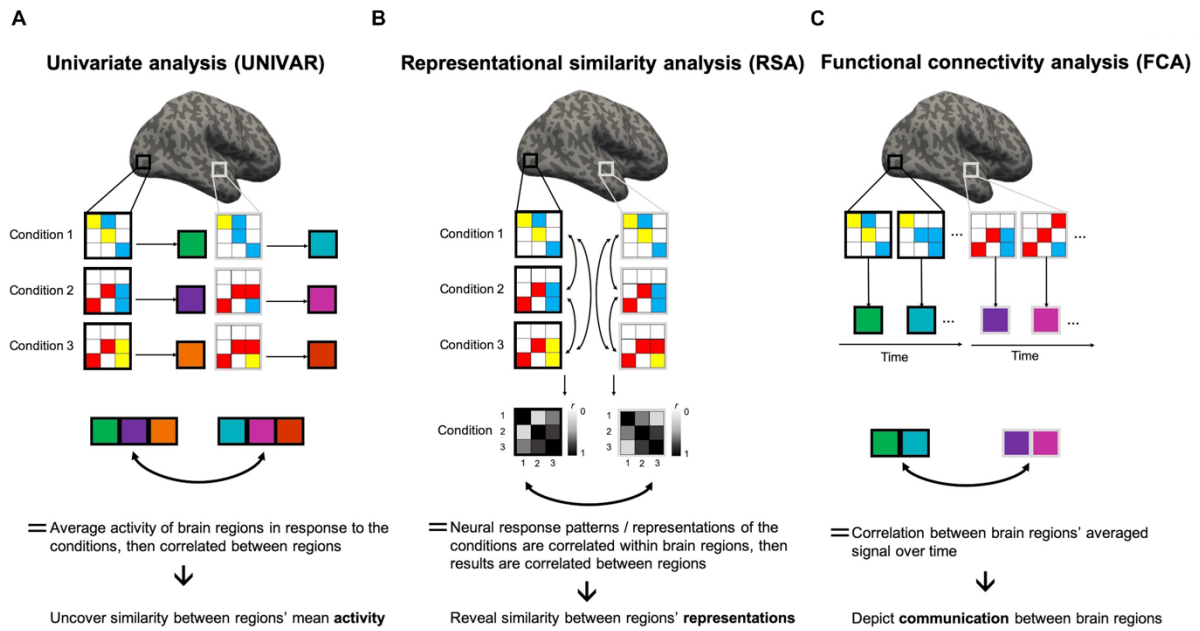


Figure 1. Representational similarity, univariate and functional connectivity analyses. Visualization of the several steps of the methodologies of these analyses to clarify the (dis)similarities. Gray and white squares drawn on the brain depict two random regions of interest (ROIs), consisting of 9 voxels each. Within these squares, colors in the 3×3 matrix depict the BOLD signal intensity of the voxels in a condition 1, 2, or 3, modeled through a general linear model (GLM). High intensity is shown in red, medium intensity in yellow, low intensity in blue. (A) In univariate analyses (UNIVAR), the signal is averaged across all voxels within the ROI, and this for every ROI and every condition. This step is visualized by large squares next to each matrix with a color depicting the calculated average. Consequently, these results are averaged across participants. To understand if the ROIs are similarly activated in light of the conditions, the results of every ROI can be correlated to each other. (B) In representational similarity analysis (RSA), the signal is not averaged across voxels within ROIs. Instead, the signal of an array of voxels in a ROI in light of a certain condition is correlated to that of another condition. All possible pairwise comparisons are made. This step is depicted by a square matrix containing the correlation results for each condition pair: the color of each element of the matrix indicates the strength of the correlation (black indicates a correlation of 1 and white indicates 0). These results show if conditions are represented similarly within a ROI. They are then averaged across subjects. Then, these results can be correlated between ROIs to understand if different ROIs represent the conditions in a similar way. (C) In functional connectivity analysis, the signal of the voxels in a ROI at different time-points (i.e., change in BOLD signal intensity over time for each voxel) is averaged across all voxels within that ROI. This step is visualized by large squares with a color depicting the calculated average. After, the resulting average BOLD signal time course can be correlated between

ROIs and consequently averaged across subjects. This analysis method depicts if two ROIs communicate.

Materials and methods

Datasets

We reanalyzed data from our previous fMRI study (Lee Masson et al., 2018). All participants provided written informed consent before the experiment in accordance with the Declaration of Helsinki. The study was carried out in accordance with the recommendations of and approved by the Medical Ethical Committee of KU Leuven (S53768 and S59577). In this study, 21 healthy participants observed grayscale videos (see Figure 2) of social touch interaction, varying in valence and arousal (Lee Masson and Op de Beeck, 2018). The experiment included 39 social touch videos, and in addition 36 non-social control videos. Participants carried out an orthogonal attention task: they pressed a button with their left or right thumb whenever the touch interaction initiator wore a gray or black shirt, depending on the instruction of that specific run (left for gray, right for black). The stimuli were displayed for 3 s, followed by an inter-stimulus interval of 3 s during which a fixation cross was presented and during which the participants could press a button as a response related to the task. Each run was divided into three blocks of 25 videos. At the start of each block, a baseline (display of a fixation cross) of 6 s was included. The total duration of each run was 7.80 min. The participants completed six runs. In the following UNIVAR and RSA analyses, we restrict the analyses to the data from the 39 social touch videos.

Importantly, when creating the videos, we controlled for the visual elements, such as clothes style and color of the actors so that these do not induce a visually biased neural response (Lee Masson and Op de Beeck, 2018). For example, having actors wear blue in the pleasant touch scenes and having actors wear red in unpleasant touch scenes can induce visual bias related to the clothing color when contrasting the brain response between the pleasant and unpleasant touch conditions.

In addition, the scan sessions included runs in which participants received (instead of observing) pleasant (brush strokes) and unpleasant touch (rubber band snaps) in a block design (see Lee Masson et al., 2018, for more details). These data were used for the intrinsic FCA.

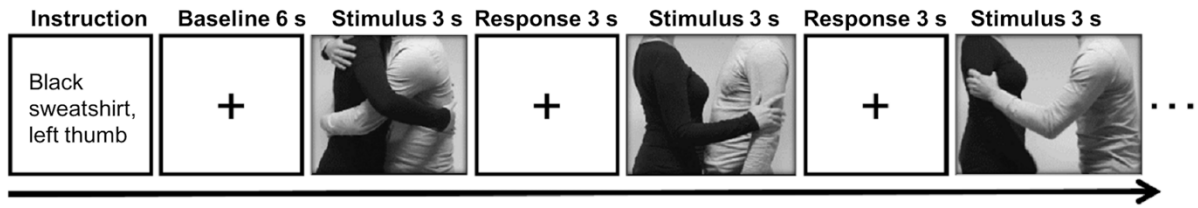


Figure 2. The experimental procedure. Participants received instruction on when they should press a certain button (e.g., press the button with your left thumb when touch interaction initiator wears black sweatshirt). After a baseline of 6 s, the stimuli were presented for 3 s always followed by an inter-stimulus interval of 3 s, during which a fixation cross was presented and participants could press a button. In this example, still frames of three social touch videos are shown (left: hug, middle: stroke, right: shake). All videos can be found here: <https://osf.io/nq5mf/>.

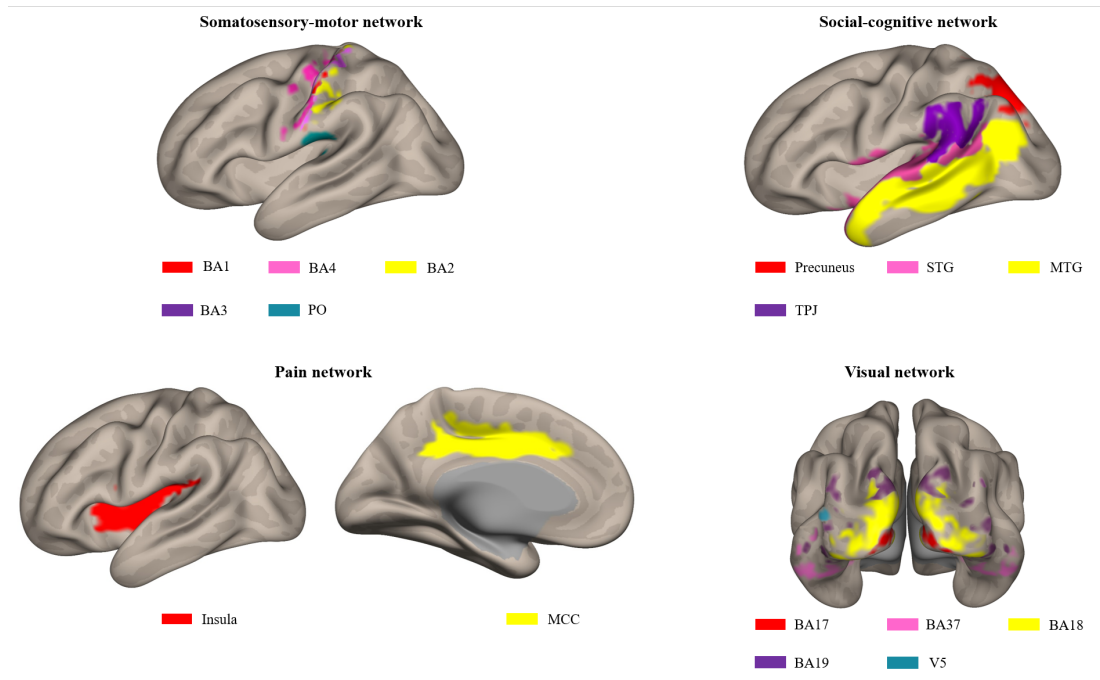


Figure 3. Illustration of the different ROIs in the context of the network they were assigned to a priori. Top left: somatosensory-motor network ROIs including BA1 (red), BA4 (pink), BA2 (yellow), BA3 (purple), and PO (blue). Top right: social-cognitive network ROIs including precuneus (red), STG (pink), MTG (yellow), and TPJ (purple). Bottom left: pain network ROIs including insula (red) and MCC (yellow). Bottom right: visual network ROIs including BA17 (red), BA37 (pink), BA18 (yellow), BA19 (purple), and V5 (blue). This figure was made using CONN toolbox 17 (Whitfield-Gabrieli and Nieto-Castanon, 2012).

Regions of interest (ROIs)

For our previous study (Lee Masson et al., 2018), we selected 16 a priori defined ROIs, belonging to four different networks in the brain that proved to be important in processing observed social touch

interactions: the somatosensory-motor network [the parietal operculum (PO), Brodmann area (BA) 3, BA1, BA2, BA4 (Rolls et al., 2003)], the social-cognitive network [the middle temporal gyrus (MTG), the precuneus, the superior temporal gyrus (STG), the temporoparietal junction (TPJ) (Jacoby et al., 2016)], the pain network [the middle cingulate cortex (MCC), the insula (Gordon et al., 2013; Lamm and Majdandžić, 2015; Morrison et al., 2011)], and the visual network given that visual stimuli were used (BA17, BA18, BA19, BA37, V5) (see Figure 3). To define these ROIs anatomically, first, we made masks with various templates from PickAtlas software (Maldjian et al., 2003, RRID:SCR_007378), SPM Anatomy toolbox (Eickhoff et al., 2005, RRID:SCR_013273) and connectivity-based parcellation atlas (Mars et al., 2012). Second, we extracted all the voxels in the mask per ROI and combined left and right hemispheres. Afterward, we examined if there were overlapping voxels among ROIs (e.g., V5 is located in BA19 and BA37) and removed overlapping voxels from each other in order to ensure all ROIs are anatomically independent. For further information about these ROIs and how they were defined, see our previous study (Lee Masson et al., 2018). In contrast to FCA that often includes a more extensive set of ROIs, the RSA method requires ROIs to contain meaningful neural signals associated with the experimental conditions. For this reason, only the aforementioned 16 ROIs, whose spatial neural patterns passed the MVPA reliability test, were selected (Lee Masson et al., 2018). Briefly, in this reliability test, runs are split into two halves and the correlation between neural patterns for within- and between-conditions are compared per ROI. This process is repeated 100 times (to randomly split the runs into two halves) and these results are then averaged. ROIs are only included if the correlations for within-condition comparisons are significantly stronger than those for between-condition comparisons. Neural pattern similarity between different conditions most likely only reflects noise when neural pattern similarity between the same conditions is low (Ritchie et al., 2017; for more details on this test and the results; see Lee Masson et al., 2018).

Univariate analysis

In our previous study, we processed functional data by using a standard preprocessing pipeline and by applying a general linear model (GLM) to each subject's data (Lee Masson et al., 2018). On top of the regressors of interest (matched to the onset time of each regressor (duration = 0) of the event-related design of the fMRI observing touch experiment), six head motion parameters were included in the models as nuisance covariates (Lee Masson et al., 2018). These GLMs were defined with data smoothed with 8 mm FWHM. For detailed information on how data was preprocessed and how the GLM was applied, see our previous study (Lee Masson et al., 2018). From these GLMs, we obtained the estimated beta-values per stimulus of the social condition (N = 39) for all voxels in each ROI. For each stimulus, we averaged the beta values of all runs, of all voxels within each ROI, and of all

participants, yielding a one-dimensional array with 39 elements in each ROI, reflecting how strongly each of the 39 videos activated the ROI. These arrays were Pearson correlated for each possible pair combination of ROIs to investigate the similarity between ROIs' average BOLD responses evoked during the observation of social touch and therefore to investigate clustering/networks of our ROIs with regard to their average activations. We refer to this clustering as the activation network.

Representational similarity analysis

This analysis was based upon a GLM applied to the fMRI observing touch experiment that consisted of 75 predictors (one for each video). The preprocessing pipeline for this analysis differed slightly from the one used for univariate analysis: data was smoothed at 5 mm FWHM. As such, we optimized the preprocessing parameters to fit the requirements of each analysis (Hendriks et al., 2017). For each ROI, we created a 39×39 neural matrix by correlating (Pearson) the multi-voxel patterns between all possible combinations of pairs of stimuli of the social condition ($N = 39$) and then averaged this across subjects (first-order RSA, Kriegeskorte et al., 2008). After, we vectorized the upper diagonal elements of this group-averaged matrix while discarding the diagonal and lower diagonal elements, and correlated (Pearson) these vectors for all possible combinations of ROI pairs (second-order RSA, Kriegeskorte et al., 2008). These comparisons between areas allow us to investigate the representational similarity between ROIs and therefore to investigate clustering/networks of our ROIs with regard to the between-condition similarity in multi-voxel activation patterns (Kriegeskorte et al., 2008). We refer to this clustering as the representation network. More information on the details of how MVPA was applied to fMRI data can be found in our previous study (Lee Masson et al., 2018).

Functional connectivity analysis

Functional connectivity analysis, performed in the CONN toolbox 17 (Whitfield-Gabrieli and Nieto-Castanon, 2012, RRID:SCR_009550), was applied to a different set of fMRI data (wherein participants received touch) obtained in the same scan sessions. We used two independent sets of fMRI data to avoid a spurious correlation between the two sets of brain networks resulting from UNIVAR and FCA. BOLD signal fluctuations may be partially induced by the presented stimuli, which may result in shared signals between networks derived from the UNIVAR and FCA methods. In the other direction, spontaneous fluctuations over time may affect the estimated univariate activation.

Preprocessing was conducted as described in our previous study (Lee Masson et al., 2018), but again optimized to fit the requirements of FCA: no smoothing was carried out to avoid a spillover effect (Alakörkkö et al., 2017). The outlier scans were detected based on the global signal spike and motion

in the functional data by the Artifact Detection Toolbox (ART) software package¹. Consequently, standard denoising methods were applied to remove confounding effects. This step consists of (1) linearly regressing out 13 principal components of white matter and cerebrospinal fluid signals, six head motion parameters and their first-order derivatives, all scrubbing covariates from the artifact detection, and main task effects (rest condition, see below), (2) linear detrending, and (3) band-pass filtering (0.008–0.09 Hz) that removes slowly fluctuating noise, such as scanner drift, and the task-induced signal. To calculate intrinsic FC (functional connectivity), we did not encode task-related information in the experimental design. Instead, task effect (i.e., receiving touch) was removed from the fMRI time series by including regressors corresponding to each task condition during the denoising step, and the rest condition was defined (Fair et al., 2007). Previous studies indicated that the intrinsic fluctuations in a BOLD signal would only be weakly affected by task demands and could be separated when entangled with the task-related signals (Fox et al., 2006; Fair et al., 2007). Several studies have implemented this approach on task-based fMRI data to yield the intrinsic functional connectivity network (e.g., Fair et al., 2007; Bassett et al., 2011; Boets et al., 2013; Ebisch et al., 2013).

For each subject, a GLM was performed to assess bivariate Pearson correlation coefficients between ROIs' BOLD time-series. These coefficients were averaged across subjects. As a result, networks of functionally connected (communicating) regions were uncovered. We refer to this clustering as the connectivity network.

Signal-to-Noise Ratio Measurement

To measure the reliability of the fMRI signal for the activation (from UNIVAR), representation (from RSA) and connectivity (from FCA) network, we randomly split the participants into two groups ($n = 10$ or 11 per group). For each of these analyses, we correlated the resulting activation, representation, and connectivity network matrices (off-diagonal values) of one group with the matrix of the other group. This process was performed for a total of 100 iterations (each time randomly splitting the data into two groups). The correlations were adjusted with the Spearman-Brown split-half reliability formula and then averaged (across the 100 iterations) for UNIVAR, RSA, and FCA separately. The results from the between-subject correlations work as a measure of signal-to-noise ratio (SNR), taking the between-subject variability in the neural data into account, in that it estimates the maximum correlation we could expect. The correlation between the same types of data from the two sub-groups (group 1 vs. group 2 in FCA results) should be higher than the correlation with another

¹ www.nitrc.org/projects/artifact_detect/

type of data (e.g., FCA vs. RSA results). This SNR correlation coefficient was also squared to obtain the proportion of the variance in the signal that can be explained by other variables.

Anatomical proximity

For each ROI per hemisphere, we collected the x-y-z coordinates of its voxels. Consequently, for each ROI pair, we calculated Euclidean distances for all possible pairs of voxels between these two ROIs. Among these calculated distances, we use the minimum value per ROI pair as a measure of the anatomical distance between the two ROIs. Then, we averaged the distances across the two hemispheres. We also performed supplementary analyses with distance based on the average rather than the minimum value, which yielded very similar results (the two indices correlate strongly, correlation $r = 0.81$). As a final step, we inverted these results to have a measure of anatomical proximity instead of distance with the maximal distance becoming the minimal proximity zero. We refer to these results as the anatomical proximity network. Dependency of functional connectivity on anatomical distance has been observed (Salvador et al., 2005). Thus, the anatomical proximity network was included in the partial correlation and the multiple regression model to rule out the effects of anatomical proximity when comparing the activation, representation and connectivity network.

Comparing the activation, representation and connectivity network

(Partial) correlation models

To understand how similar the activation, representation, connectivity and anatomical proximity network are, we conducted a rank-order correlational analysis between these networks. In addition, we also computed the partial Spearman correlation coefficient to understand the similarities between the two networks while controlling for the remaining networks. To draw statistical inferences, we conducted a permutation test, wherein one of the variables of interest [one of the networks, consisting of all possible unique ROI pairs (120 pairs)] was randomly shuffled and then (partially) correlated with the unshuffled variables [remaining original networks, each consisting of all possible unique ROI pairs (120 pairs per network)]. This process was iterated a 1000 times. These permutation tests provide empirical p-values (probability values) reflecting the proportion of permutations wherein the (partial) correlations with the shuffled data were larger (or equally large) than the original (partial) correlations.

Multiple regression models

Following up on the (partial) correlation models, we conducted multiple regression analysis to investigate if the activation, representation or connectivity network respectively, could be explained

by the other remaining networks. The anatomical proximity network was also included in all of the multiple regression models. Z-score standardizations were performed to normalize the data before building a regression equation. Similarly to the correlational analysis, permutation tests were used to obtain empirical p-values. In the end, the percentage variance explained by the model was compared to the squared signal-to-noise ratio of the predicted variables of the model.

Multi-dimensional scaling (MDS) and Procrustes transformations

We conducted multi-dimensional scaling (MDS) on the activation, representation, and connectivity network matrices to visualize the networks in a two-dimensional space that shows the distance between each pair of ROIs based on how dissimilar these ROIs are in terms of their activation, representation, and connectivity respectively. We used the built-in MATLAB function “mdscale” with default parameters, minimizing the default goodness-of-fit criterion “stress” and 100 replicates of the scaling. MDS results of the representation network were used as a template to which the MDS results of the activation and connectivity networks were aligned using Procrustes transformations, to visualize the networks in the same space. For this, we used the built-in MATLAB function “procrustes” with default parameters.

Results

Networks

In total, we have four matrices (see Figure 4). For three of the methods (RSA, UNIVAR, and FCA) the values in the matrices are based upon correlational analyses. In each of these matrices, we had a large range of values. For the representation network, for which vectorized first-order RSA results were correlated between all ROI pairs, correlations range from 0.07 (V5 – insula) to 0.82 (BA3 – BA4). In the activation network matrix, the correlation results range from -0.01 (precuneus – PO) to 0.98 (BA3–BA4). The values of the ROI-to-ROI connectivity range from -0.17 (precuneus – PO) to 0.83 (BA3–BA4). The anatomical proximity network values range from 0 to 67.63. The higher the value, the more closely the two ROIs are located. As the values are inverted distances, a value of 0 indicates the minimum anatomical proximity between ROIs (e.g., BA1–BA17), which in the original distance was 67.63 mm. A proximity value of 67.63 indicates the maximum anatomical proximity between ROIs: these ROIs are located right next to each other (e.g., BA1–BA2).

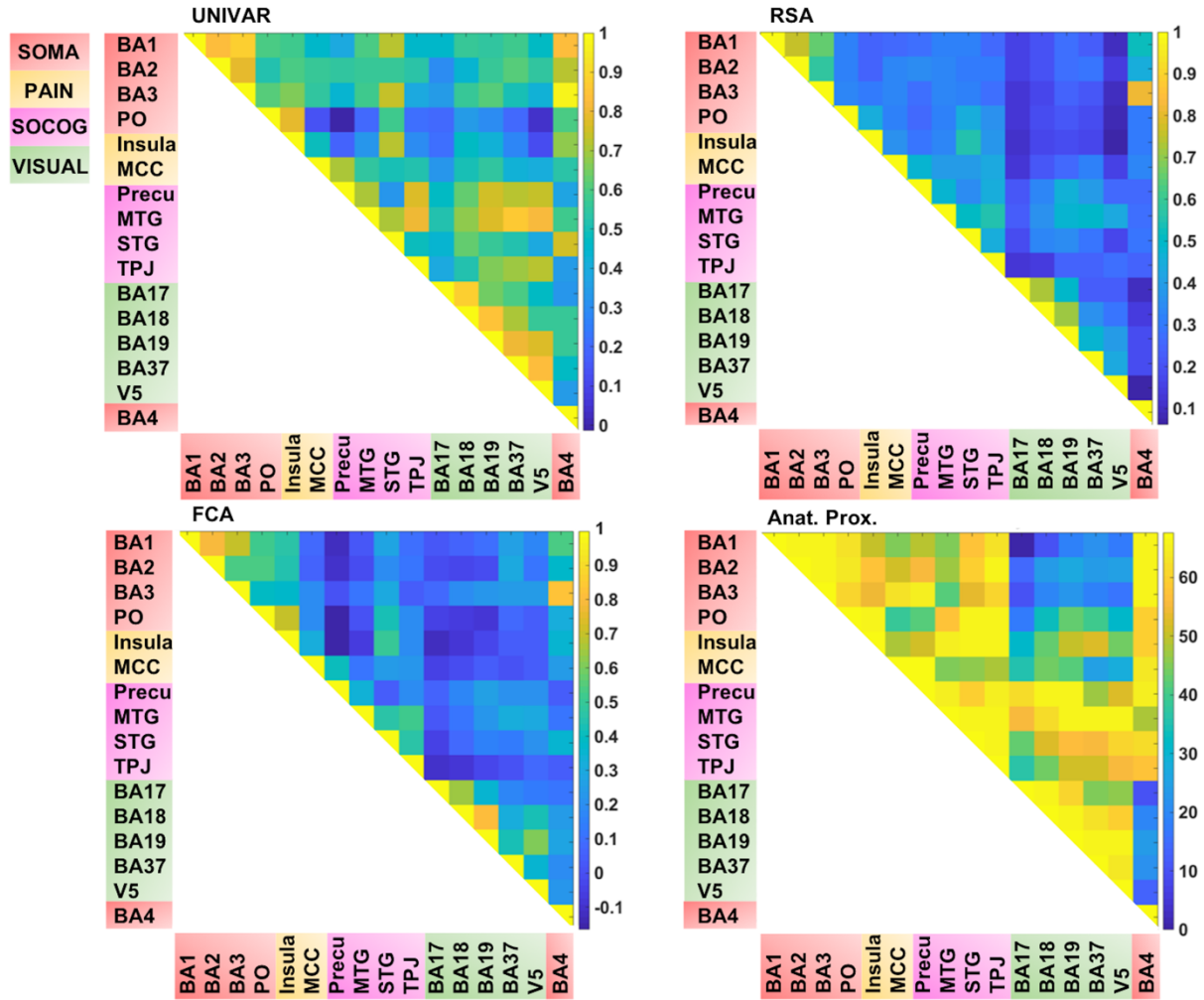


Figure 4. Visualization of the different networks before z-score standardization of the correlation coefficients. Top left: activation network (UNIVAR), top right: representation network (RSA), bottom left: connectivity network (FCA), bottom right: anatomical proximity network. Yellow in the matrix = when two ROIs are very similar in their activation (UNIVAR results) or their representation (RSA results), are well connected (FCA results), or are located closely in the brain. Blue in the matrix = when two ROIs are very different in their activation (UNIVAR results) or their representation (RSA results), are not connected (FCA results), or are located remotely in the brain. SOMA (red), somatosensory-motor network areas; PAIN (yellow), pain network areas; SOCOG (purple), social-cognitive network areas; VISUAL (green), visual network areas.

Each of the matrices was very reliable. The signal-to-noise ratio estimated from the results of between-subjects correlations was $r = 0.92$ (squared to obtain explainable variance: $EV = 85\%$) for the representation network, $r = 0.96$ ($EV = 92\%$) for the activation network, and $r = 0.97$ ($EV = 94\%$) for the connectivity network. As illustrated in Figure 4, the representation, the activation, and connectivity networks look highly similar to each other. For example, the high correlation values between the ROIs

in somatosensory areas such as BA3, BA1, and BA2 are apparent in all of these networks: these ROIs contain similar task-related information (based on RSA results), are activated to a similar level (based on UNIVAR results) and are functionally linked to each other (based on FCA results). BA4, the motor area, is strongly correlated to BA3 and BA1 in the representation, activation, and connectivity network, but only moderately to BA2. Another example is the moderate to high correlation between visual areas, found in the representation, activation and connectivity network. In sum, this finding applies to all four ROI networks. Areas of different ROI networks typically show lower correlations, which is again consistent across methods. For example, the moderate correlations between social-cognitive areas and visual areas can be found in the representation, activation and connectivity network.

Comparing networks

To understand the (dis)similarity between representation, activation, connectivity, and anatomical proximity networks more quantitatively, we tested the linear relationship among these networks. The results indicated that all networks are similarly organized in the context of brain function and anatomy, with the Spearman rank-order correlations (all significant) ranging from 0.53 to 0.79 (see Table 1). In addition, the partial correlation coefficients were computed between two networks after removing the effect of the other remaining networks. The results from partial correlation (including all four networks) demonstrated that, after controlling for the other networks, the activation and connectivity network (Spearman correlation $r_s = 0.50$, $p < 0.001$), the representation and activation network ($r_s = 0.34$, $p < 0.001$), and the representation and anatomical proximity network ($r_s = 0.67$, $p < 0.001$) still correlate significantly (see Table 2). Conversely, the measured partial correlation between the representation and connectivity network was no longer significant after ruling out the effects of the other covariates (partial $r_s = 0.10$), implying that their association is fully explained by their relationship with other networks. The partial correlation between the connectivity and anatomical proximity network (partial $r_s = 0.15$), and between the activation and anatomical proximity network (partial $r_s = 0.06$) was also no longer significant.

	UNIVAR	RSA	FCA	Anat. Prox.
UNIVAR	1	0.66*	0.70*	0.53*
RSA	0.66*	1	0.61*	0.79*
FCA	0.70*	0.61*	1	0.55*
Anat. Prox.	0.53*	0.79*	0.55*	1

Table 1. The correlations between the activation (UNIVAR), representation (RSA), connectivity (FCA) and anatomical proximity (Anat. Prox.) network. * *Significant correlation.*

	UNIVAR	RSA	FCA	Anat. Prox.
UNIVAR	1	0.34*	0.50*	−0.10
RSA	0.34*	1	0.10	0.67*
FCA	0.50*	0.10	1	0.15
Anat. Prox.	−0.10	0.67*	0.15	1

Table 2. The partial correlations between the activation (UNIVAR), representation (RSA), connectivity (FCA) and anatomical proximity (Anat. Prox.) network.

As an alternative approach, we also implemented multiple regression models. Similar to the (partial) correlation measurements, these regression models quantify the relations between the networks, but in addition the regression models provide an estimate of the total variance in a network that can be explained by all other networks.

A first model tested if the connectivity, activation and anatomical proximity network significantly predicted the *representation* network. The coefficient of determination from the regression equation indicated that these three predictors explained 71.2% of variability in the representation network [$R^2 = 0.712$, $F(3, 116) = 96$, $p < 0.001$]. The squared signal- to-noise ratio (based on the between-subjects correlation) in the representation network indicated 85% of the variance to be explainable, leaving approximately 14% of the signal unexplained. In addition, we calculated the β coefficients to examine the degree to which each predictor independently contributes to the prediction of the representation network. According to the results, the anatomical proximity network significantly contributed to the prediction of the representation network ($\beta = 0.40$, $p < 0.001$), as did the connectivity network ($\beta = 0.36$ $p = 0.004$) and the activation network ($\beta = 0.26$, $p = 0.03$).

Similarly, we predicted the *connectivity* network based on the representation, activation, and anatomical proximity network, using multiple regression analysis. The results indicated that the predictors explained 59.6% of variability in the connectivity network [$R^2 = 0.596$, $F(3, 116) = 57$, $p < 0.001$]. The squared signal-to-noise ratio (based on the between-subjects correlation) in the connectivity network indicated 94% of the variance to be explainable, leaving approximately 34% of the signal unexplained. When examining the independent contributions of each predictor, we found out that the representation network significantly contributed to the prediction of the connectivity network ($\beta = 0.51$, $p = 0.003$), as did the activation network ($\beta = 0.37$, $p = 0.005$), but not the anatomical proximity network ($\beta = -0.07$, $p = 0.55$).

Lastly, we tested if the representation, connectivity, and anatomical proximity network significantly predicted the *activation* network. The results revealed that the predictors explained 55.8% of variability in the activation network [$R^2 = 0.558$, $F(3, 116) = 49$, $p < 0.001$]. The squared signal-to-noise ratio (based on the between-subjects correlation) in the activation network indicated 92% of the variance to be explainable, leaving approximately 36% of the signal unexplained. The predictors indicated that the representation network significantly contributed to the prediction of the activation network ($\beta = 0.41$, $p = 0.01$), as did the connectivity network ($\beta = 0.41$, $p = 0.001$), but not the anatomical proximity network ($\beta = -0.01$, $p = 0.91$).

Thus, for each type of network, we find that a lot of the organization can be predicted from the other networks, but there is also some remaining variance left unexplained. We visualized this unique signal left in each of these networks after regressing out the signal explained by the other networks from the representation, activation and connectivity network respectively (see Figure 5B). In Figure 5B, in contrast to Figures 4, 5A (which takes the values of Figure 4 and z-score standardizes them, for reasons mentioned above), the networks now do not look similar: they show different patterns.

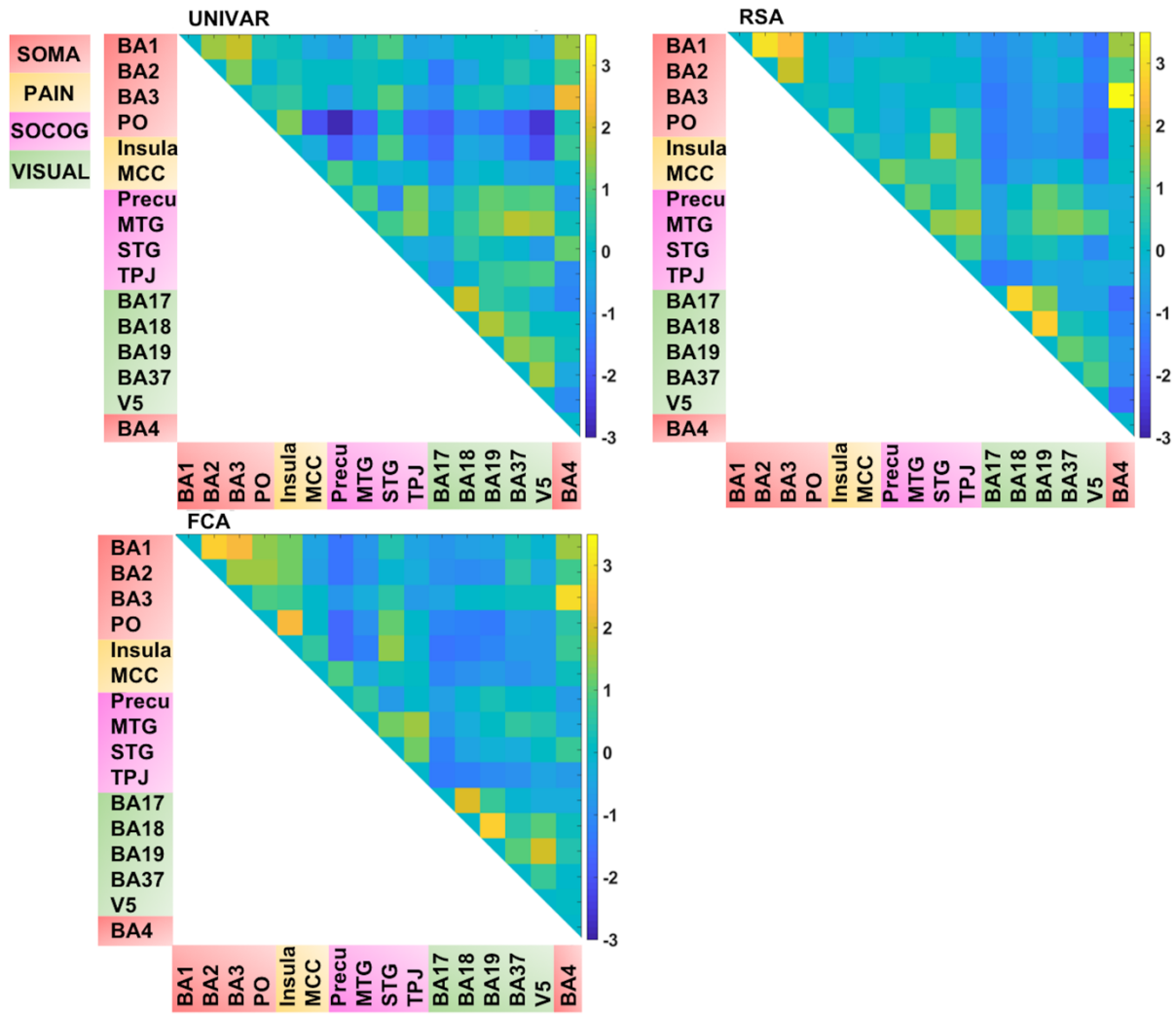


Figure 5A. Visualization of the different networks after z-score standardization of the correlation coefficients. Top left: activation network (UNIVAR), top right: representation network (RSA), bottom left: connectivity network (FCA), bottom right: anatomical proximity network. Yellow in the matrix = when two ROIs are very similar in their activation (UNIVAR results) or their representation (RSA results), are well connected (FCA results), or are located closely in the brain. Blue in the matrix = when two ROIs are very different in their activation (UNIVAR results) or their representation (RSA results), are not connected (FCA results), or are located remotely in the brain. SOMA (red), somatosensory-motor network areas; PAIN (yellow), pain network areas; SOCOG (purple), social-cognitive network areas; VISUAL (green), visual network areas.

Several unique findings concerning correlations between ROI-networks can be observed in Figure 5B. For example, social- cognitive brain areas correlate strongly to other visual areas in the activation network (e.g., r (before z-score standardization) = 0.69 between TPJ and BA37) while this is moderate to low in the representation (e.g., r = 0.24 between TPJ and BA37) and connectivity network (e.g., r = 0.01 between TPJ and BA37). This finding implies that these areas are activated similarly, but do not

represent similar information nor do they communicate with each other. Another example, social-cognitive areas correlate moderately to somatosensory-motor areas (e.g., $r = 0.32$ (representation), $r = 0.59$ (activation) between MTG and BA1), except in the connectivity network (e.g., $r = 0.03$ between MTG and BA1). As the last example, visual area V5 shows a moderate correlation to other brain areas in the representation network (e.g., $r = 0.39$ between V5 and BA19) while a much stronger correlation is found in the other networks (e.g., $r = 0.76$ (activation) $r = 0.61$ (connectivity) between V5 and BA19).

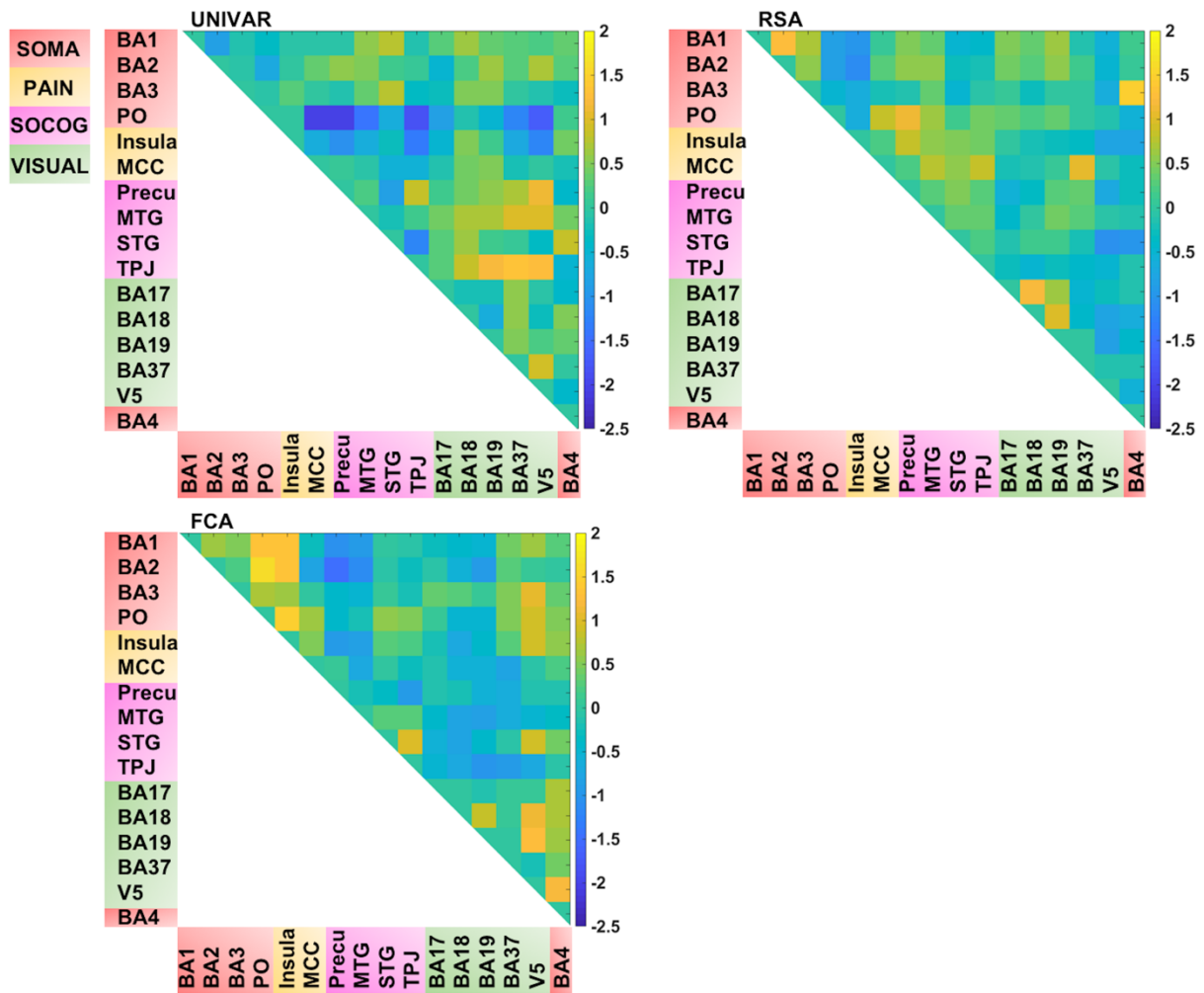


Figure 5B. Visualization of the different networks after regressing out the signal explained by the other networks. Top left: activation network (UNIVAR), top right: representation network (RSA), bottom left: connectivity network (FCA), bottom right: anatomical proximity network. Yellow in the matrix = when two ROIs are very similar in their activation (UNIVAR results) or their representation (RSA results), are well connected (FCA results), or are located closely in the brain. Blue in the matrix = when two ROIs are very different in their activation (UNIVAR results) or their representation (RSA results), are not connected (FCA results), or are located remotely in the brain. SOMA (red), somatosensory-motor network areas; PAIN (yellow), pain network areas; SOCOG (purple), social-cognitive network areas; VISUAL (green), visual network areas.

For visualization purposes, we performed MDS on the three types of dissimilarity matrices to reconstruct two-dimensional spatial configuration that reflects the proximity in the matrices. Moreover, Procrustes transformations were performed to align the configurations. The resulting configurations are shown in Figure 6. The results confirm the high similarity (d (Procrustes distance: the difference between the shape of the two networks) between the activation and representation

network = 0.48, d between the connectivity and representation network = 0.34, d between the activation and connectivity network = 0.42) and some dissimilarities between the networks as was previously indicated by the (partial) correlation and multiple regression models. As an example of correspondence between the three networks, Figure 6 shows that somatosensory-motor areas are located nearby in all three networks, implying high similarity in activation and representation and strong inter- regional communication among these areas. As an example of a difference between the networks, the social-cognitive brain areas are placed close to visual areas overall in the activation network (blue in Figure 6) but not so much in the other networks. It suggests that social cognitive brain areas and the visual cortex do not represent the same information and that those areas are not functionally connected despite the similar magnitude of neural response. Other visualization methods such as principal component analysis result in similar plots as Figure 6.

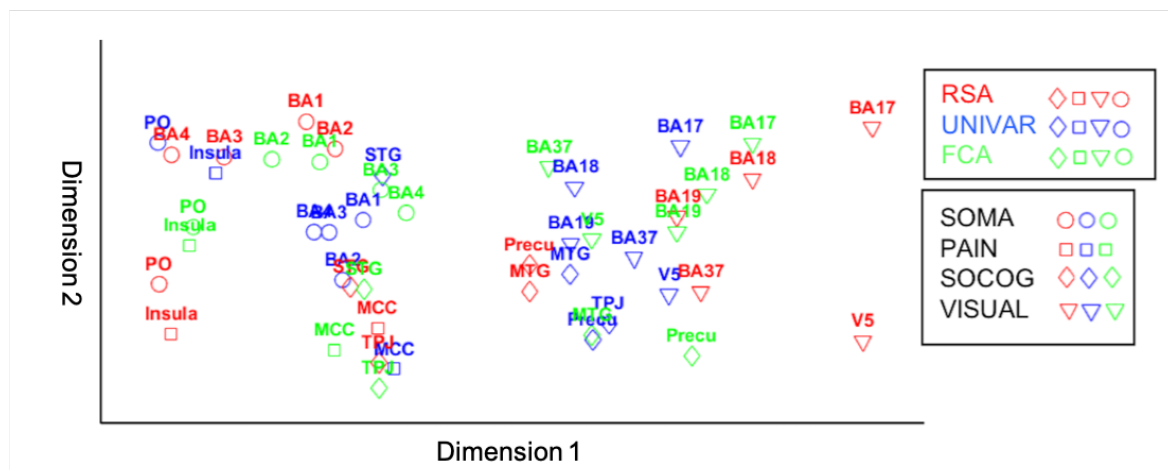


Figure 6. Procrustes transformed MDS results of the activation (UNIVAR, blue) and connectivity network (FCA, green) to the MDS results of the representation network (RSA, red). SOMA (circles), somatosensory-motor network areas; PAIN (squares), pain network areas; SOCOG (diamonds), social-cognitive network areas; VISUAL (triangles), visual network areas.

Summary

The representation, the activation, and connectivity networks were proven to be reliable based on their signal-to-noise ratio. All three functional networks are organized highly similarly. For example, moderate correlations between social-cognitive areas and visual areas can be found in the representation, activation and connectivity network. Based on multiple regression models, we found that for each type of network, a lot of the organization can be predicted from the other networks, but there is also some remaining variance left unexplained. For example, social- cognitive areas and somatosensory-motor areas are activated similarly (UNIVAR) and represent similar information (RSA), but do not communicate with each other (FCA). In other words, the networks derived from RSA, UNIVAR, and FCA, although highly similar, also each contain unique information. We have visualized these results using Procrustes transformations on MDS results in Figure 6.

Discussion

RSA has recently emerged as a method for investigating how brain regions are organized into networks. UNIVAR and FCA are two other popular methods for analyzing fMRI data to understand functional networks. Although two or more of these methods have been used simultaneously to analyze the same set of data in many studies, most of them have focused on the properties of each ROI separately. No study, to our knowledge, directly and simultaneously compared networks derived from RSA with networks built with UNIVAR and FCA. In the current study, we explored how the organization of networks built from RSA, UNIVAR, and FCA relate to each other after ruling out the effect of the anatomical location of network nodes (ROIs). We analyzed fMRI data of a previous study (Lee Masson et al., 2018) with these methods and performed (partial) correlation and multiple regression analysis on the resulting networks.

The current study reveals that neural networks resulting from RSA, UNIVAR, and FCA are highly similar even after ruling out the effect of anatomical proximity. As predicted, brain areas within the somatosensory-motor network are similarly activated, represent similar task-related information, and are intrinsically connected. This also applies to the other sub-networks (pain, social-cognitive and visual). As outlined in the introduction, RSA, UNIVAR, and FCA share theoretical and/or methodological properties that can explain similarities as observed in this study. The high similarity in the neural networks of RSA and FCA provides support for the idea that brain areas showing similar stimulus-related selectivity are also intrinsically connected. Our finding is in line with previous resting-state fMRI studies that have identified functionally relevant networks, such as the primary visual network,

auditory network, motor network, and cognitive networks, during rest (e.g., Biswal et al., 1995; Fox and Raichle, 2007; Jung et al., 2018).

On the other hand, our finding suggests that the network derived from each method contains unique signals. To reveal this, we compared the explainable variance of each network revealed by SNR estimation with the actual variance explained by the other networks. These results suggested that the network, derived from each method, contains unique information that none of the other networks are able to explain. Analyzing the remaining signal variance that was left unexplained, we were also able to reveal unique network organizations of each method. For example, brain areas in the social-cognitive network are similar to areas in the visual network in terms of neural activation, whereas neural patterns of those two sub-networks do not represent the same information and they are not intrinsically connected. Another example is the moderate correlations between social-cognitive brain areas and somatosensory-motor areas in the activation and representation network, but not in the connectivity network.

This unique information is important to keep in mind when interpreting a functional network found with one particular method. Although second-order RSA can be used to construct brain connectivity, RSA and FCA adopt different approaches shown in their methodology: correlating RDMS in RSA; correlating the BOLD signal fluctuations in FCA. Thus, RSA is used for investigating the similarity between brain areas in how they represent task-related information while FCA is used for investigating how a series of brain areas construct the intrinsically connected cortical network. These distinctions allow RSA and FCA to tap into the functional architecture of the brain from different perspectives as revealed in the unique information the networks contain.

Likewise, the same reasoning can be applied to the relationship between UNIVAR and FCA, and RSA and UNIVAR. As outlined in the introduction, they are related theoretically and empirically while they differ in their focus, allowing both similarities and dissimilarities between the resulting networks.

Such distinctions between networks derived from different methods have also been observed in the recent study of Jung and her colleagues (Jung et al., 2018) comparing resting-state fMRI and structural connectivity. Although their comparison involves different methods than ours, they provided some possible explanations that should be considered in the current study. The quality and nature of the datasets used for three methods (even from identical data sources, but measured at different times or analyzed in a different way) may not be equal and different measurement noise may be present (Jung et al., 2018). In addition, they mention that networks during mental activity are modulated away (slightly) from intrinsic connections, which is especially relevant to the comparison of RSA with FCA. Accordingly, our findings of similarities and differences between RSA and FCA networks are consistent with the observation that studies using both RSA and FCA lead to either similar or different conclusions

about brain function derived from the two methods (e.g., Boets et al., 2013; Zeharia et al., 2015). Several other factors could play a role in the (dis)similarities between the networks derived from RSA, FCA, and UNIVAR, for example vasculature. Despite the high similarity across the networks derived from the UNIVAR, RSA, and FCA methods, given the nature of uniqueness of each network, we encourage researchers to understand the benefits of each methodology and what they (do not) detect; and to use them adequately depending on the research questions.

As a critical note, we point to several limitations of our current study. First, the current findings are based on only one task domain (i.e., social touch scene perception), and our conclusions should be complemented by future studies that include other tasks, such as moral decision-making tasks, or tasks using other sensory modalities such as auditory and tactile scenes. RSA and UNIVAR methods may not produce similar network organizations in another task. This could in particular be the case, when having a task with no activation differences across the conditions but evoking neural pattern selectivity. We hypothesize that it will be a general phenomenon; that also in other domains there will be a shared network organization that dominates with, in addition, a smaller unique component. Second, we selected a limited number of ROIs rather than including a large number of network nodes. One important argument for doing this is that the selected brain regions had to include meaningful task-related signals for performing RSA (see the description of diagonal vs. non-diagonal measures as a reliability test in choosing ROIs in Methods). The effect of the number and size of ROIs on the relationships between the networks obtained using RSA, UNIVAR, and FCA can be explored further. Finally, extending the comparisons made in the current study is another important step to take. Specifically, networks built from second-order RSA and multivariate functional connectivity could also be compared (Coutanche and Thompson-Schill, 2014; Anzellotti and Coutanche, 2018). Other analyses than RSA exist that can pick up on more complex spatial relations between ROIs (e.g., Haxby et al., 2011; Ozay et al., 2012; Firat et al., 2013; Onal et al., 2017). Although RSA is one of the most commonly used approaches in cognitive neuroscience in comparing neural spaces across different brain regions, uncovering the complexity of the spatial relationships captured through more advanced approaches may help to compare the three methodologies in depth.

Conclusion

The present study provides first-time evidence that cortical networks derived from three commonly used neuroimaging approaches (RSA, UNIVAR, and FCA) are highly similar regardless of the structural variations of each network. Importantly, the study also demonstrates that each of these three networks contains unique information, unexplainable by the other networks. As such, all three methods are important when investigating the functional organization of networks in the brain.

Improving the understanding of the relationship between the functional networks derived from these methods will allow researchers to use RSA, UNIVAR, and FCA more adequately.

Data availability statement

The data for this study can be found in the Open Science Framework (<https://osf.io/b4np9/>).

Ethics statement

All participants provided written informed consent before the experiment in accordance with the Declaration of Helsinki. The study was carried out in accordance with the recommendations of and approved by the Medical Ethical Committee of KU Leuven (S53768 and S59577).

Author contributions

IP, HO, and HL set up the project, fMRI data collection (see Lee Masson et al., 2018), and wrote the manuscript. IP and HL analyzed the data.

Funding

This work was supported by the European Research Council (Grant No. ERC-2011-Stg-284101), a federal research action (Grant No. IUAP-P7/11), Fonds Wetenschappelijk Onderzoek Vlaanderen (Grant No. G088216N), a Hercules Grant (Grant No. ZW11_10), and an Excellence of Science grant (Grant No. G0E8718N) to HO, and Marguerite-Marie Delacroix Foundation (GV/B-336) to HL.

Acknowledgements

This manuscript has been released as a preprint at BioRxiv (doi: 10.1101/487199).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Alakörkkö, T., Saarimäki, H., Glerean, E., Saramäki, J., and Korhonen, O. (2017). Effects of spatial smoothing on functional brain networks. *Eur. J. Neurosci.* 46, 2471–2480. doi: 10.1111/ejn.13717

- Anzellotti, S., and Coutanche, M. N. (2018). Beyond functional connectivity: investigating networks of multivariate representations. *Trends Cogn. Sci.* 22, 258–269. doi: 10.1016/j.tics.2017.12.002
- Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., and Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *PNAS* 108, 7641–7646. doi: 10.1073/pnas.101898 5108
- Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409
- Boets, B., Op De Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science* 342, 1251–1254. doi: 10.1126/science.1244333
- Bulthé, J., Prinsen, J., Vanderauwera, J., Duyck, S., Daniels, N., Gillebert, C. R., et al. (2018). Multi-method brain imaging reveals impaired representations of number as well as altered connectivity in adults with dyscalculia. *Neuroimage* 190, 289–302. doi: 10.1016/j.neuroimage.2018.06.012
- Cordes, D., Haughton, V. M., Arfanakis, K., Carew, J. D., Turski, P. A., Moritz, C. H., et al. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in “resting-state” data. *AJNR Am. J. Neuroradiol.* 22, 1326–1333.
- Cordes, D., Haughton, V. M., Arfanakis, K., Wendt, G. J., Turski, P. A., Moritz, C. H., et al. (2000). Mapping functionally related regions of brain with functional connectivity MR imaging. *AJNR Am. J. Neuroradiol.* 21, 1636–1644.
- Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us? *Cogn. Affect. Behav. Neurosci.* 13, 667–673. doi: 10.3758/s13415-013-0186-2
- Coutanche, M. N., and Thompson-Schill, S. L. (2014). Using informational connectivity to measure the synchronous emergence of fMRI multi-voxel information across time. *J. Vis. Exp.* 89:e51226. doi: 10.3791/51226
- Crossley, N. A., Mechelli, A., Vértes, P. E., Winton-Brown, T. T., Patel, A. X., Ginestet, C. E., et al. (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. *PNAS* 110, 11583–11588. doi: 10.1073/pnas.1220826110
- Davis, T., LaRocque, K. F., Mumford, J. A., Norman, K. A., Wagner, A. D., and Poldrack, R. A. (2014). What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *Neuroimage* 97, 271–283. doi: 10.1016/j.neuroimage.2014. 04.037

- Ebisch, S. J. H., Mantini, D., Romanelli, R., Tommasi, M., Perrucci, M. G., Romani, G. L., et al. (2013). Long-range functional interactions of anterior insula and medial frontal cortex are differently modulated by visuospatial and inductive reasoning tasks. *Neuroimage* 78, 426–438. doi: 10.1016/j.neuroimage.2013.04.058
- Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034
- Fair, D. A., Schlaggar, B. L., Cohen, A. L., Miezin, F. M., Dosenbach, N. U. F., Wenger, K. K., et al. (2007). A method for using blocked and event-related fMRI data to study “resting state” functional connectivity. *Neuroimage* 35, 396–405. doi: 10.1016/j.neuroimage.2006.11.051
- Firat, O., Ozay, M., Onal, I., Öztekin, I., and Yarman-Vural, F. T. (2013). Functional mesh learning for pattern analysis of cognitive processes. *ICCI-CC 2013*, 161–167. doi: 10.1109/ICCI-CC.2013.6622239
- Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., and Raichle, M. E. (2006). Spontaneous neural activity distinguishes human dorsal and ventral attention systems. *PNAS* 103, 10046–10051. doi: 10.1073/pnas.0604187103
- Fox, M. D., and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711. doi: 10.1038/nrn2201
- Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Hum. Brain Mapp.* 2, 56–78. doi: 10.1002/hbm.460020107
- Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008
- Geerligs, L., Henson, R. N., and Cam-Can. (2016). Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation. *Neuroimage* 135, 16–31. doi: 10.1016/j.neuroimage.2016.04.047
- Gilron, R., Rosenblatt, J., Koyejo, O., Poldrack, R. A., and Mukamel, R. (2017). What’s in a pattern? Examining the type of signal multivariate analysis uncovers at the group level. *Neuroimage* 146, 113–120. doi: 10.1016/j.neuroimage.2016.11.019
- Gordon, I., Voos, A. C., Bennett, R. H., Bolling, D. Z., Pelphrey, K. A., and Kaiser, M. D. (2013). Brain mechanisms for processing affective touch. *Hum. Brain Mapp.* 34, 914–922. doi: 10.1002/hbm.21480
- Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: the early beginnings. *Neuroimage* 62, 852–855. doi: 10.1016/j.neuroimage.2012.03.016

- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430. doi: 10.1126/science.106 3736
- Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., et al. (2011). A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* 72, 404–416. doi: 10.1016/j.neuron.2011.08.026
- Haynes, J. D. (2015). A primer on pattern-based approaches to fMRI: principles, pitfalls, and perspectives. *Neuron* 87, 257–270. doi: 10.1016/j.neuron.2015.05. 025
- Haynes, J. D., and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534. doi: 10.1038/nrn1931
- Hendriks, M. H. A., Daniels, N., Pegado, F., and Op de Beeck, H. P. (2017). The effect of spatial smoothing on representational similarity in a simple motor paradigm. *Front. Neurol.* 8:222. doi: 10.3389/fneur.2017.00222
- Jacoby, N., Bruneau, E., Koster-Hale, J., and Saxe, R. (2016). Localizing pain matrix and theory of mind networks with both verbal and non-verbal stimuli. *Neuroimage* 126, 39–48. doi: 10.1016/j.neuroimage.2015.11.025
- Jimura, K., and Poldrack, R. A. (2012). Analyses of regional-average activation and multivoxel pattern information tell complementary stories. *Neuropsychologia* 50, 544–552. doi: 10.1016/j.neuropsychologia.2011.11.007
- Jung, J., Visser, M., Binney, R. J., and Lambon Ralph, M. A. (2018). Establishing the cognitive signature of human brain networks derived from structural and functional connectivity. *Brain Struct. Funct.* 223, 4023–4038. doi: 10.1007/ s00429- 018- 1734- x
- Kaas, J. H. (1997). Topographic maps are fundamental to sensory processing. *Brain Res. Bull.* 44, 107–112. doi: 10.1016/S0361-9230(97)00094-4
- Kriegeskorte, N. (2011). Pattern-information analysis: from stimulus decoding to computational-model testing. *Neuroimage* 56, 411–421. doi: 10.1016/j. neuroimage.2011.01.061
- Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* 2:4. doi: 10.3389/neuro.06.004.2008
- Lamm, C., and Majdandžić, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy - A critical comment. *Neurosci. Res.* 90, 15–24. doi: 10.1016/j.neures.2014.10.008
- Lee Masson, H., and Op de Beeck, H. (2018). Socio-affective touch expression database. *PLoS One* 13:e0190921. doi: 10.1371/journal.pone.0190921

- Lee Masson, H., Van De Plas, S., Daniels, N., and Op de Beeck, H. (2018). The multidimensional representational space of observed socio-affective touch experiences. *Neuroimage* 175, 297–314. doi: 10.1016/j.neuroimage.2018.04.007
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869–878. doi: 10.1038/nature06976
- Maldjian, J. A., Laurienti, P. J., Kraft, R. A., and Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19, 1233–1239. doi: 10.1016/S1053-8119(03)00169-1
- Mars, R. B., Sallet, J., Schüffegen, U., Jbabdi, S., Toni, I., and Rushworth, M. F. S. (2012). Connectivity-based subdivisions of the human right “temporoparietal junction area”: evidence for different areas participating in different cortical networks. *Cereb. Cortex* 22, 1894–1903. doi: 10.1093/cercor/bhr268
- Morrison, I., Björnsdotter, M., and Olausson, H. (2011). Vicarious responses to social touch in posterior insular cortex are tuned to pleasant caressing speeds. *J. Neurosci.* 31, 9554–9562. doi: 10.1523/JNEUROSCI.0397-11.2011
- Mur, M., Bandettini, P. A., and Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI - An introductory guide. *Soc. Cogn. Affect. Neurosci.* 4, 101–109. doi: 10.1093/scan/nsn044
- Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430. doi: 10.1016/j.tics.2006.07.005
- Onal, I., Ozay, M., Mizrak, E., Öztekin, I., and Yarman-Vural, F. T. (2017). A new representation of fMRI signal by a set of local meshes for brain decoding. *IEEE Trans. Signal. Inform. Process. Over Netw.* 3, 683–694. doi: 10.1109/TSIPN.2017.2679491
- Ozay, M., Öztekin, I., Öztekin, U., and Yarman-Vural, F. T. (2012). Mesh learning for classifying cognitive processes. *arXiv*. [preprint]. doi: 10.1109/EMBC.2013.6611113
- Raizada, R. D. S., and Kriegeskorte, N. (2010). Pattern-information fMRI: new questions which it opens up and challenges which face it. *Int. J. Imaging Syst. Technol.* 20, 31–41. doi: 10.1002/ima.20225
- Raizada, R. D. S., Tsao, F. M., Liu, H. M., and Kuhl, P. K. (2010). Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: prediction of individual differences. *Cereb. Cortex* 20, 1–12. doi: 10.1093/cercor/bhp076
- Ritchie, J. B., Bracci, S., and Op de Beeck, H. (2017). Avoiding illusory effects in representational similarity analysis: what (not) to do with the diagonal. *Neuroimage* 14, 197–200. doi: 10.1016/j.neuroimage.2016.12.079

- Rolls, E. T., O'Doherty, J., Kringelbach, M. L., Francis, S., Bowtell, R., and McGlone, F. (2003). Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cereb. Cortex* 13, 308–317. doi: 10.1093/cercor/13.3.308
- Salvador, R., Suckling, J., Coleman, M. R., Pickard, J. D., Menon, D., and Bullmore, E. (2005). Neurophysiological architecture of functional magnetic resonance images of human brain. *Cereb. Cortex* 15, 1332–1342. doi: 10.1093/cercor/ bhi016
- Whitfield-Gabrieli, S., and Nieto-Castanon, A. (2012). CONN: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* 2, 125–141. doi: 10.1089/brain.2012.0073
- Xue, S. W., Weng, X. C., He, S., and Li, D. W. (2013). Similarity representation of pattern-information fMRI. *Chin. Sci. Bull.* 58, 1236–1242. doi: 10.1007/s11434- 013- 5743- 0
- Yang, Z., Fang, F., and Weng, X. (2012). Recent developments in multivariate pattern analysis for functional MRI. *Neurosci. Bull.* 28, 399–408. doi: 10.1007/ s12264- 012- 1253- 3
- Zeharia, N., Hertz, U., Flash, T., and Amedi, A. (2015). New whole-body sensory-motor gradients revealed using phase-locked analysis and verified using multivoxel pattern analysis and functional connectivity. *J. Neurosci.* 35, 2845–2859. doi: 10.1523/JNEUROSCI.4246-14.20

Chapter 2: The position of visual word forms in the anatomical and representational space of visual categories in occipitotemporal cortex

This chapter as provided in this thesis is currently under review with the journal *Imaging Neuroscience*.

Abstract

Recent reviews emphasized the need for a study at the individual level at a high spatial resolution that investigates the complexity of multiple subareas of word selectivity and how this relates to selectivity for other visual categories. First, we mapped category selectivity (determined using a general linear model) on the brain surface of each participant. Second, given that many dimensions organize object space (such as animacy and object size), it is important to investigate word selectivity in the representational space formed by a large set of categories. To achieve these goals, we presented 19 participants with images of 20 different categories during 7T fMRI. In the left hemisphere, we found three subareas of the VWFA and one extra subarea around the pFus face-selective area. We also found several areas of selectivity to hands that could consistently guide the localization of word and face areas. Results of the right hemisphere were less clear. A competition between words and human-related categories could not be found within the representational space. These results allow a consistent and reliable way to locate subareas of word selectivity and may inspire future research on the competition between words and other categories.

Introduction

Visual object recognition is supported by lateral and ventral parts of the occipitotemporal cortex (OTC) (Barton et al., 2002; Grill-Spector et al., 2000; Moutoussis et al., 2002; Gaillard et al., 2006). This region of cortex contains areas that show a selectivity towards certain visual categories, like faces (Kanwisher et al., 1997; Tsao et al., 2006), bodies (Downing et al., 2001; Peelen & Downing, 2005), hands (Bracci et al., 2010; Bratch et al., 2018), tools (Bracci et al., 2012; Peelen et al., 2013), scenes (Aguirre et al., 1998; Epstein & Kanwisher, 1998), word forms (Cohen et al., 2000; Cohen et al., 2002; Baker et al., 2007), and numerals (Shum et al., 2013; Abboud et al., 2015). In addition to such local category-selective areas, other categories are represented through response patterns distributed throughout the OTC (Haxby et al., 2001; Kriegeskorte et al., 2008). These different patterns of activity for different categories can be detected using a decoding approach called multi-voxel pattern analysis (MVPA).

An important characteristic of the functional organization of OTC is that category selectivity is organized according to dimensions on which categories vary. For example, in ventral OTC, areas selective to animate categories are localized more laterally than areas selective to inanimate categories (Grill-Spector & Weiner, 2014). This dimension also plays an organizing role in representational spaces as found through MVPA (for a classic example, see Kriegeskorte et al., 2008). Other examples of such dimensions that play a role in the organization of the OTC are real-world size (Konkle & Oliva, 2012) and body topography (Orlov et al., 2010). More low/mid-level visual dimensions like retinotopy also influence OTC's functional organization (Malach et al., 2002). For example, Grill-Spector and colleagues (2017) found that areas selectively responding to faces show a bias for visual information from the fovea. Areas selective to word forms have a similar preference for the foveal part of the visual field (Hasson et al., 2002; Rauschecker et al., 2012; Le et al., 2017; Lerma-Usabiaga et al., 2021).

Previous research found that some category-selective areas appear in a predictable location based on anatomical landmarks and relative to other areas of selectivity. For example, Weiner and colleagues (2014) found that the anterior tip of the mid-fusiform sulcus, that separates the fusiform gyrus into a lateral and medial part, predicts the location of fusiform face area 2 (FFA-2), also called pFus (posterior fusiform face area). As another example, Weiner and Grill-Spector (2010) uncovered that face- and limb-selective areas alternate in a consistent pattern. This enabled a reliable way to define face areas (mFus, pFus and IOG) based on their spatial relation to limb-selective areas (along OTS and ITG) (Weiner & Grill-Spector, 2010). This alternation of category-selective areas did appear to be more consistent across subjects in the right hemisphere as compared to the left hemisphere (Weiner & Grill-Spector, 2010). The authors proposed that this may be due to the presence of word-

selective areas in the left hemisphere (Weiner & Grill-Spector, 2010), although they could not test this because of the absence of a word form category in their study.

Weiner's and Grill-Spector's (2010) proposal, that the face-limb area alternation is not consistent in the left hemisphere due to words, is grounded in research that proposes a competition for cortical territory in OTC between faces and words, based on the neural recycling theory (Dehaene & Cohen, 2007; Dehaene et al., 2010). This theory posits that, during development of recently invented skills (such as written language), circuits that are used for similar evolutionary older skills, are repurposed for this new skill (Dehaene & Cohen, 2007). The specifics of how this recycling unfolds is a topic of debate, as there are studies that indicated a destructive competition (Dehaene et al., 2010), but also studies that indicated there is no competition but rather a fine-tuning of visual object recognition (Hervais-Adelman et al., 2019). A recent study adds another category on top of faces and words to the mix: Nordt and colleagues (2021) found that selectivity for faces and words increases while learning to read and write, whereas selectivity to limbs decreases. This could indicate both a destructive competition and a visual fine-tuning. This result also questions the specificity of the competition during development of reading and writing: other categories than faces and words could very well play a role.

The word selectivity found in ventral OTC is often called the visual word form area (VWFA). This area of selectivity extends from the posterior occipitotemporal sulcus (OTS) to about the midpoint of the fusiform gyrus (Yeatman et al., 2021). Two recent studies have attempted to subdivide this large region into parts based on functional differences. Lerma-Usabiaga and colleagues (2018) applied two types of contrasts (perceptual and lexical) to discriminate between parts of the long strip of activation that forms the VWFA. They found that, compared to the posterior region (pOTS), the anterior region (mOTS) was more sensitive to lexical characteristics of the stimuli and was structurally more connected to language areas. In their discussion, they also link these subdivisions to previous research, stating that the mOTS corresponds to the central VWFA (as described in Vogel et al., 2012; Dehaene et al., 2002; Cohen et al., 2002; Cohen et al., 2004) and pOTS to posterior VWFA (as described in Vogel et al., 2012; Ben-Shachar et al., 2007). White and colleagues (2019) investigated selective spatial attention of the VWFA by presenting subjects with two words at once. They concluded that VWFA-1 could in parallel process multiple words, whereas VWFA-2 processed, after an integration of hemifields, only one word at a time. VWFA-1 and VWFA-2 correspond to pOTS and mOTS. Yeatman and colleagues (2021) in their review propose that VWFA-1 lies next to FFA-1 and VWFA-2 next to FFA-2 and in between those, a body/limb-selective area (see also Grill-Spector & Weiner, 2014). These studies divided the large VWFA into two subareas based on functional properties, but not based on clear anatomical boundaries between those subareas.

Several researchers in the field have recently emphasized the need for studies that investigate the location of the VWFA and possible subareas, with high anatomical precision at the level of the individual brain (without spatial normalization or smoothing) (Caffarra et al., 2021; Yeatman et al., 2021). Such single-subject precision can have a great impact on a study's ability to accurately describe spatial organization: it is possible that when the VWFA seems like one large region (with a gradually increasing sensitivity to lexical characteristics from posterior to anterior), this could be due to a lack of such precision (Caffarra et al., 2021). With a higher precision, the VWFA may be subdivided based on anatomical boundaries between them, instead of (only) based on functional differences. Thus, in our study, we mapped word selectivity (obtained through a general linear model) on the anatomy of the ventral brain surface of each of the 19 subjects that were scanned with 7T fMRI. This allowed a high spatial resolution, investigating each subject separately without applying normalization or smoothing preprocessing steps. This type of study will allow for consistency in the definition of subareas of the VWFA based on anatomical boundaries between them and can consequently allow (more) studies into the functional distinctions between these subareas (Caffarra et al., 2021; Yeatman et al., 2021).

On top of the words category, we included many other potentially relevant categories in our study. This allows word selectivity to be related to other category selectivity. Such a relation may be expected based on for example Weiner and Grill-Spector (2010) who found such a consistent spatial relation between face- and limb-selective areas (but they did not include word-selective areas). Based on the neural recycling theories, we expect the categories of faces, hands, and bodies to be obvious candidates to relate to the functional neuroanatomy of word selectivity. Also based on Yeatman and colleagues (2021) we expected bodies and hands (a type of limb) to be relevant because, like described above, it might lie between two subregions of the VWFA. One recent study by Boring and colleagues (2021) did look at faces compared to words at a high spatial resolution in the individual subject space, but no other categories were included. Thus, to achieve this goal, we put several other relevant visual categories (faces, hands, bodies, and other character categories: fake script and numbers) on the word selectivity map and related them. In addition, we ran a split-half analysis, to replicate the found category selectivity and understand better the response profile of these areas of selectivity. Uncovering a spatial relation between word and other category-selective areas could allow for more reliability in future studies when locating category-selective areas.

Given the multiple dimensions that organize object space (such as animacy, object size, and retinotopy), another contribution of our study, that includes many categories that vary on these dimensions, is that it allows the use of MVPA to investigate word selectivity in the representational space. The competition between words and faces (and in extension possibly also categories such as hands and bodies) proposed by the neural recycling theories can then be investigated in this

representational space. To create such a representational space, we included in our study a total of 20 categories that varied on several dimensions such as animacy and object size.

Methods

Subjects

Nineteen subjects participated in the study (mean age: 30.1 ± 6.8 (23-45), sex: 11 males, 8 females). Sixteen subjects were right-handed, three were left-handed (subject 2, 7 and 17). All subjects had normal or corrected to normal visual acuity. Every subject gave informed consent. All procedures were approved by the ethics committee of Vrije Universiteit Amsterdam and adhered to the guidelines of the Declaration of Helsinki.

Stimuli

Stimuli were presented on a 32" LCD screen (69.8×39.3 cm, 120 Hz) designed to use in an MRI environment (BOLDscreen, Cambridge Research Systems, UK). The screen resolution was 1920×1080 pixels. The screen was positioned at the end of the bore and viewed through a mirror (distance from screen: 220 cm) mounted on the head coil. All stimuli were presented with MATLAB (MathWorks, Inc.) and the Psychophysics Toolbox Version 3 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). Stimuli were presented at a semi-random position around the middle of the screen in every trial. A fixation dot was always present in the middle of the screen, presented over the stimuli.

We chose 20 conditions (19 categories + 1 scrambled control condition) so that we would have a rich and varied set according to several criteria. The set included both natural and artificial objects, known and unknown shapes (cubies and smoothies, see Op de Beeck et al., 2006), animate and inanimate shapes, objects with a different real-world size, objects differing in how they are used and whether they are a tool or not, ... The categories could be grouped together in several ways according to dimensions such as these. Within each category, different viewpoints/angles on the objects were included and there was sufficient variability in the identity of the stimuli. The following is an alphabetical list of all the categories, for each an example is also depicted in Figure 1: bodies, buildings, cars, cats, chairs, cubies, faces, fake script, fish, flowers, hammers, hands, musical instruments, numbers, scissors, scrambled, smoothies, trees, vegetables, words.

In the creation of the stimuli, we attempted to minimize the role of low- and mid-level features in the differences between categories. For eight (bodies, buildings, cars, cats, chairs, faces, fish, hammers) of the 20 categories, stimuli were provided to us by Michael Cohen (Cohen et al., 2017). They ensured high within-category variability by for example, images of items in different angles, in different positions, ... (Cohen et al., 2017). As described above, we also followed these guidelines in the creation of the other categories' images as well. The following categories (partially) consist of images from the Bank of Standardized Stimuli (BOSS): musical instruments, flowers, vegetables,

scissors (Brodeur et al., 2010). To add more images to these categories, and to create stimuli for other categories (hammers, hands, trees), we created stimuli ourselves by using freely available images on the internet.

The control category (scrambled images) was created using the `randblock(rgb)` function, that is available online for MATLAB (Mathworks, Inc.), on all other images in the stimulus set. From all these scrambled images, we chose 30 images (including at least one from each of the other 19 conditions) where the middle of the image contained more variability than the borders as this resembles the general way in which the other categories' images were organized.

For the unknown objects we used the cubies and smoothies stimuli described in the study by Op de Beeck and colleagues (2006). Both objects could vary on four different shape-dimensions. On each of the dimensions, the object had a value from 0 to 5. We chose 16 shapes for both cubies and smoothies that took an extreme position in this 4D-space (e.g., 0505, 0005) and four more that had a center position within this space (e.g., 2222, 2323). Ten of those 20 stimuli were duplicated before any further processing happened. All 30 stimuli were then flipped with a random angle, making sure the duplicates were angled differently than their original version, to end up with 30 unique stimuli. During creation, the objects' positions on the grey background were ensured to not always be in the center.

The words, the fake script and the numbers categories consisted of two, three or four letters/symbols. The letter/symbol strings in these three categories were angled ascending or descending to have a difference in aspect ratio between stimuli. Furthermore, the letters/symbols were shown in a bold font filled with a random dotted pattern. Half of the stimuli of each category were presented onto a white background and the other half on a black background. These properties assured that these conditions would be more comparable to the other categories in terms of retinotopic envelope. The words did not have any meaning in the lexicon, thus no semantic representation. For simplicity, we refer to this category as words rather than letter strings. These words stimuli all contained at least 1 vowel (except one of the 30 stimuli) and were all pronounceable. To create the fake script stimuli, a combination of two different fonts was used to ensure the letters would take on a varied shape (e.g., short, long, blocked, curvy).

Several conditions required extra control as to match low and mid-level features between categories, on top of ensuring within-condition variability in viewpoints and identity. Hammers and scissors are often elongated in shape. To avoid retinotopic confounds, we translated and rotated exemplars, included different viewpoints and selected the scissors open instead of closed. We controlled for the specific shape of trees by using different kinds of trees, by sizing them smaller/larger and by not always having them appear in the middle part of the background but also on the upper/lower left/right part. We replaced the images of the buildings category that contained trees

with other images of buildings that did not contain trees, as trees is supposed to be a separate category. The bodies' images were also adapted so they would have no visible hands, as this is a category on its own.

In the end, there were 30 stimuli available for each category. All these stimuli were then matched for average luminance, contrast, and spectral energy by the SHINE toolbox (Willenbockel et al., 2010), inspired by the study of Cohen and colleagues (2017). The results of the toolbox were checked in that the identity of each stimulus was ensured to be at least moderately clear/visible to viewers. We conducted several additional checks: a mean greyvalue image, a standard deviation greyvalue image and an edge image was created for each category and then compared between categories to ensure they did not differ regarding these criteria.

All images are available in the following OSF project: <https://osf.io/y8vnf/>.

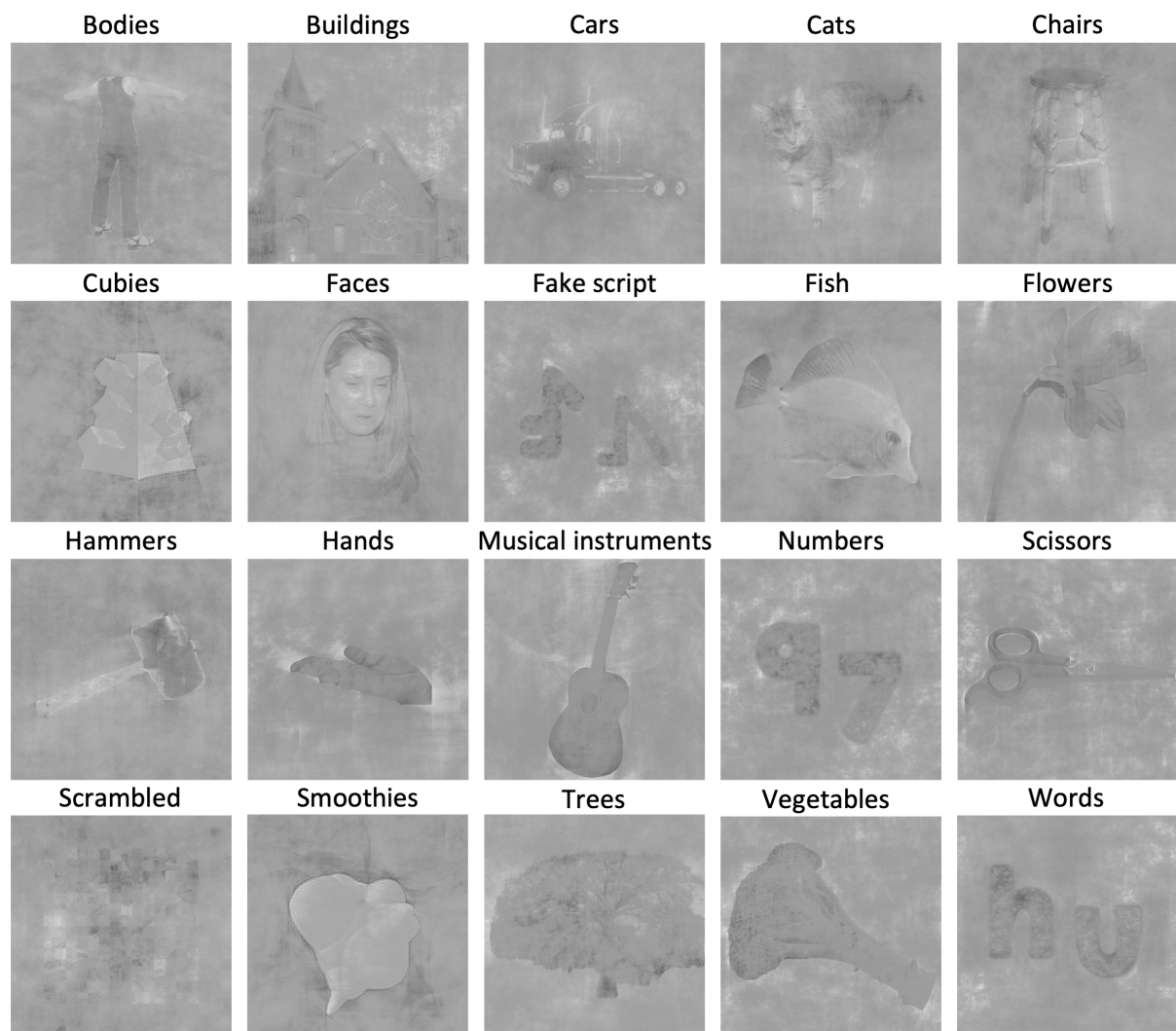


Figure 1. For each of the 20 categories, an example stimulus is shown. The category the stimulus belongs to is indicated on top of the image.

Experimental design

Subjects were explained the task before entering the scanner. They were asked to focus on the fixation dot in the center of the screen. A dual task was used. The first task was a one-back task. The subjects pressed a button with either their left or right hand (this instruction changed every subject) when an image was an identical copy of the image shown in the previous trial. Thus, if the same object was presented in successive trials, but rotated or angled differently, they did not press the button as this was not an identical copy of the image. For the second task, subjects pressed a button as soon as possible as a change in category occurred. They used the button in the other hand than the hand used for the first task. This second task was meant to ensure subjects were processing category information and not just focusing on certain more low-level features to perform the first task. For this reason, after explaining the tasks, the subjects were introduced to all the categories of the experiment by showing them an example stimulus of each category.

The subjects completed seven runs (in one subject only six, sometimes eight) inside the scanner. Within a run, we included 40 blocks that each lasted 10.05 seconds consisted of 15 trials/stimuli, and 3 blocks (at the start, middle and end of the run) that contained only a fixation dot and lasted 15 seconds. A run thus lasted 447 seconds. After the first fixation block, in the first half of the run, 20 blocks were presented, one for each category, in a random order. After the middle fixation block, in the second half of the run, the 20 blocks appeared in the reversed order that was used in the first half of that run. A trial lasted 0.67 seconds, for 40% of the trial (0.27 seconds) only the fixation dot was present, for the other 60% of the trial (0.40 seconds) the stimulus was present. The presentation order of the stimuli within a block was always random and two repeats of stimuli (for the task) were programmed at random points within a block. We thus ended up with 140.7 seconds of data for each category in each subject with 7 runs. There were two types of runs: the 15 images for each block/category were either from the first (run type 1) or the second half (run type 2) of the full set of 30 images.

(f)MRI acquisition

Data was acquired with a 7T Philips Achieva MRI scanner (Philips Healthcare, Best, The Netherlands) with an 8 channel transmit coil and a 32 channel receive coil (Nova Medical Inc, Wilmington, United States) at the Spinoza Centre for Neuroimaging in Amsterdam (the Netherlands). Functional and anatomical imaging happened while making use of universal pulses (for more explanation see Gras et al., 2017). During brain imaging, respiratory data was collected using a respiration belt and cardiac data was collected using a pulse oximeter on a finger of the left hand.

We used a 3D-EPI sequence. The relevant sequence parameters are volume repetition time (TR) = 1.37 s, echo time (TE) = 16.9 ms, flip angle = 13°, voxel size = 1.79x1.79x1.8 mm, field-of-view (FOV) = 200x200x176 mm and matrix size = 112 x 112 x 98. We also collected images with an identical sequence except for reversed phase-encoding blips or in other words, phase-encoding in the opposite direction. Those images are used to correct for distortions. For all subjects an anatomical scan was collected using a MPRAGE sequence with parameters: TR = 10 ms, TE = 3.3 ms, flip angle = 8°, spatial resolution = 0.8x0.8x0.8 mm, matrix size = 288x288x205.

Preprocessing

(f)MRI data

The dataset was formatted according to BIDS (Gorgolewski et al., 2016) after converting the PAR/REC files to nifti files using DCM2NIIx (Li et al., 2016) and reorienting them to RAS+ with nibabel in Python. Preprocessing steps were carried out using fMRIPrep 20.2.0 (Esteban et al., 2018a; Esteban et al., 2018b; RRID:SCR_016216), which is based on Nipype 1.5.1 (Gorgolewski et al., 2011; Gorgolewski et al., 2018; RRID:SCR_002502). In short, the 7T BOLD images underwent susceptibility distortion correction, realignment, coregistration to the T1 weighted image and spatial normalization. In addition, surface reconstruction was also performed based on the anatomical image with Freesurfer, as implemented in fMRIPrep. In our analyses, we did not use the spatially normalized version of our data, as we wanted to remain within the individual space of every subject's brain to contain as much spatial and anatomical detail as possible. FMRIprep provided details on all the preprocessing in its report and this can be found below. In addition, we looked at tSNR images (average of the timeseries divided by the standard deviation) for every subject to check the quality of the fMRI data.

In subjects 4 until 10, because of a temporary problem with the coil, a small part of the right hemisphere was darker than the left hemisphere. For the anatomical data, a shadow like this would influence the surface reconstruction. To avoid this, we performed several steps. For each anatomical scan, we checked the distribution of intensity values using ITK-SNAP (Yushkevich et al., 2006) and then clipped them off accordingly; the values were then rescaled using ANTs (Tustison et al., 2014); denoising and bias correction was applied to the scans using SPM12 and cat12. Surface reconstruction results (from a fMRIPrep anat only process) were carefully checked and improved. All subjects' brain surfaces were, in this process, successfully reconstructed. After this, using these results, we ran fMRIPrep in full to also perform functional preprocessing steps.

This shadow did not present trouble for the functional data preprocessing. Preprocessing results were carefully checked through fMRIPrep reports and through checking if several contrasts (e.g. faces versus all other conditions) were of similar strength and size with (1) a general linear model based on

the functional space after realignment, versus (2) a general linear model after realignment, susceptibility distortion correction and coregistration to T1w space, versus (3) a general linear model after all these steps plus normalization to MNI space. This latter check was meant to ensure all spatial preprocessing worked as expected.

What follows is the detailed description provided by fMRIPrep. “A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.3.3 (Avants et al. 2008; RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823; Zhang et al., 2001). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847; Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438; Klein et al., 2017). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c [Fonov et al., 2009; RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym]. For each of the six/seven/eight (dependent on the subject) BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. A B0-nonuniformity map (or fieldmap) was estimated based on two (or more) echo-planar imaging (EPI) references with opposing phase-encoding directions, with 3dQwarp Cox and Hyde (1997) (AFNI 20160207). Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbrregister (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9; Jenkinson et al., 2002). The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): fsaverage, fsnative. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space

by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152Nlin2009cAsym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al., 2014) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al., 2002). FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by Power et al., 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor; Behzadi et al., 2007). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's aseg segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize

the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer). Many internal operations of fMRIPrep use Nilearn 0.6.2 (Abraham et al., 2014; RRID:SCR_001362), mostly within the functional processing workflow.” For more details of the pipeline, see the section corresponding to workflows in fMRIPrep’s documentation.

Physiological data

The physiological data was formatted according to BIDS (Gorgolewski et al., 2016) using `scanphyslog2bids` created by Lukas Snoek and available through GitHub (<https://github.com/lukassnoek/scanphyslog2bids>). Correction for physiological noise was performed with RETROICOR (Glover et al., 2000; Hutton et al., 2011) using Fourier expansions of different order for the estimated phases of cardiac pulsation (3rd order), respiration (4th order) and cardio-respiratory interactions (1st order) (Harvey et al., 2008). The corresponding confound regressors were created using the MATLAB PhysIO toolbox (Kasper et al., 2017), open-source code available as part of the TAPAS software collection (<https://www.translationalneuromodeling.org/tapas>). The physiological data collection was not successful for all participants: subjects 2, 3, 4, 5 and the first functional run of subject 18 did not include physiological data.

Data analyses

Univariate analysis

The BOLD response of each subject, run and voxel was modeled using a general linear model (GLM, also called univariate analysis) with the Statistical Parametric Mapping toolbox (SPM12, Wellcome Centre for Neuroimaging, London) in MATLAB (Mathworks inc.). Within this model, each condition was captured by a regressor that was set with certain onsets within the run and with a certain duration. This yielded 21 regressors (20 categories + one fixation condition) per run. In addition, several confound regressors were included, obtained from either fMRIPrep preprocessing or from the PhysIO toolbox. Specifically, the realignment parameters and their temporal derivatives (total: 12) were included to minimize motion effects on the data. For each subject and run, fMRIPrep also provided motion outliers regressors (a binary regressor for each volume labeled as outlier). These were also included in the GLM. The number of outliers differed for every run and subject. Whenever both respiratory and cardiac data were available for a subject, the resulting confound regressors (total: 18) were included into the model. Each condition’s results were then contrasted against all other conditions (excluding the fixation condition) to show the voxels that respond more to that condition than to all other conditions, to create a map of the functional neuroanatomy in response to all our conditions, in each of our subjects. In the visualization of this functional neuroanatomy, we focused on

the following categories (contrasting is versus all other categories except fixation): faces, bodies, hands, words, numbers, fake script, and a typical objects category (chairs) to provide a reference landmark. This will be depicted by a figure where the result of these contrasts is shown on the reconstructed brain surface of every subject. Contrasts are family-wise error corrected and significance is set at $p < 0.05$.

All the analyses happen within the subject-specific space, thus without performing spatial normalization of the data of that subject. We followed the advice by Weiner and Grill-Spector (2013) as summarized in Table 2, that encourages this approach because the data stays true to the gyral and sulcal patterns of a specific brain, which allows a more accurate localization of activity. They also advise not to perform spatial smoothing as this can result in inaccurate localization of activity and in addition, this may result in averaging of regions that are in truth distant on the surface. We also followed the pipelines as described by Brodoehl and colleagues (2020): for each subject, the GLM analysis is performed without first spatially normalizing or smoothing the data in the volume-space. These results can then be projected onto the brain surface.

Split-half analysis

To quantify category selectivity and show its replicability, we ran a split-half analysis. First, on the left and right hemisphere surface of each subject, we drew a region of interest (ROI) that contained the middle-anterior ventral surface of the occipito-temporal cortex (see Figure 2A). This ROI was drawn so that all second and third clusters of selectivity to words, faces and hands (see Results) would be included. One subject (subject 10) did not show these clusters of activity in the right hemisphere, so no ROI could be drawn.

Second, for every subject, we ran two GLMs: one using the odd runs and one using the even runs. Using the odd runs GLM, we computed the contrasts of interest, including one condition, either faces, bodies, hands, words, numbers, fake script, or an objects category (chairs), versus all other 19 categories. These contrasts were set with an uncorrected $p < .0005$. Per subject, for each of these contrasts, we made an intersection of each functional contrast with the left and the right hemisphere anatomical ROI. As a result, 7 anatomical-functional ROIs were created per hemisphere for each subject: a ventral middle-anterior face, body, hand, word, number, fake script, and object area. It was not possible to create each of these ROIs for every subject: during the making of the intersection, there were not always voxels that reached significance.

Third, using the even runs GLM, we selected the beta values of all 21 conditions per ROI and per subject, and averaged them across all runs and all voxels within an ROI. As a result, we ended up with one beta value per category, per ROI and per subject. At last, we averaged across subjects to create a

bar graph depicting the averaged beta value per category per ROI and calculated the standard error for each category/bar.

Within each ROI, we selected the main category of relevance for the functional contrast defining the ROI, e.g., faces for the left ventral middle-anterior face area. Using a paired t-test, we compared this category to the other categories used to select ROIs, in the case of a face area this would be faces against either bodies, hands, words, fake script, numbers, and chairs. Thus, per ROI, we performed 6 paired t-test between the main category of that area and the 6 other categories of particular interest. These tests were Bonferroni corrected per ROI (6 tests), meaning that significance was reached at $p < .008$, which was also depicted in the bar graphs using a *. In addition, we performed a paired t-test on the difference between the first and second preferred category between the hand- and body-selective voxels in the left hemisphere. We also performed several paired t-tests on the difference of two categories between the left and right hemisphere ROIs: between the word areas for the difference between words and numbers and for the difference between words and fake script (Bonferroni corrected $p < .03$) and between the hand areas for the difference between hands and words.

Multi-voxel pattern analysis (MVPA)

FreeSurfer performs an automated aparc parcellation using the Desikan-Killiany atlas. From this parcellation, we constructed a large occipito-temporal cortex (OTC) region of interest (ROI) separately for each hemisphere, by including the fusiform gyrus, inferior temporal, and lateral occipital cortical regions according to the aforementioned parcellation (e.g., Mattioni et al., 2020; Lerma-Usabiaga et al., 2018, see Figure 2B). We used the CoSMoMVPA toolbox (Oosterhof et al., 2016) for MATLAB (Mathworks, Inc.). A multi-voxel pattern in response to each of the conditions for each run was constructed by using the beta coefficient estimates for all voxels present within the ROI. We used the cross-validated Mahalanobis distance (Walther et al., 2016), also called linear discriminant contrast (LDC) to decode the condition for every possible pair of conditions (excluding the fixation condition), for example, faces versus scenes, words versus numbers, The code for this was written by J. Brendan Ritchie. The results were represented in a so-called dissimilarity matrix where each point in the matrix showed the distance between the condition linked to that row and the condition linked to that column of the matrix. The distance signified a measure of dissimilarity between the multi-voxel pattern of the ROI in response to the row-condition versus the column-condition. The higher the distance, the more dissimilar the patterns of these two conditions were. The distance estimates for each condition pair were created using a cross-validation scheme where the data is split up into training and testing folds according to the standard leave-one-run-out partitioner. Each run was used as the testing fold while all others were used for the training fold once, generating as many distance estimates as there were

runs available for that subject. These estimates were then averaged, and this average was placed in the appropriate spot in the dissimilarity matrix. Then, we normalized each subject matrix, per ROI, by dividing all the values inside by the maximum value of that matrix. Consequently, we averaged the matrix across subjects per ROI. To show the reliability of these matrices, per ROI, each subject matrix was correlated with the average (calculated without that subject) matrix. We performed several paired t-tests (with a Bonferroni corrected p -value) between different distances, to confirm observations made from the multidimensional scaling plots (described below).

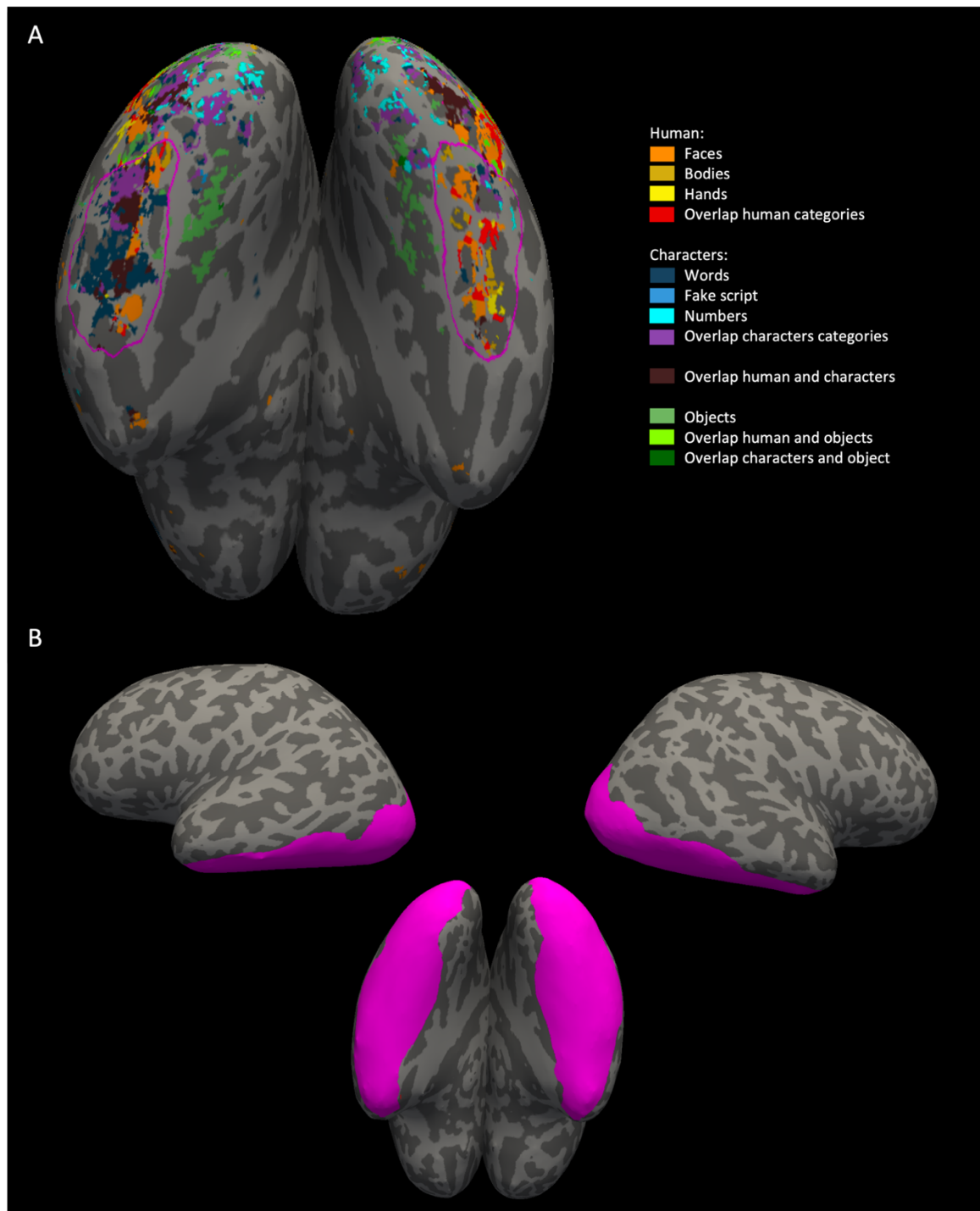


Figure 2. (A) The region of interest (ROI) for the split-half analysis for both the left and right hemisphere, shown using a pink outline on a brain surface of one of the participants. The ROI includes all middle and anterior selectivity to words (dark blue on the surface), faces (orange) and hands (yellow). (B) The region of interest (ROI) for multi-voxel pattern analysis, here shown for both the left and right hemisphere of the occipitotemporal cortex (OTC) in pink on a brain surface of one of the participants. The left and right OTC ROI was created separately by combining the lateral occipital cortex, inferior temporal gyrus, and fusiform gyrus from the Desikan-Killiany atlas incorporated in FreeSurfer.

Multidimensional scaling (MDS) and Procrustes transformations

Multidimensional scaling (MDS) was used to visualize the main dimensions underlying the patterns in the representational dissimilarity matrices in a two-dimensional space where the distance between the points in this space was a measure for how dissimilar these points were: the higher the distance, the more dissimilar. We used the built-in MATLAB (Mathworks, Inc.) function `mdscale` with the default parameters, while minimizing the default goodness-of-fit criterion: stress and 100 replicates of the scaling. We applied MDS on the matrices, averaged across subjects, keeping left and right OTC separate. In addition, we also applied MDS on the normalized dissimilarity matrices of each individual participant. The MDS results of every subject, per ROI, were transformed to the average MDS results using a Procrustes transformation (using the built-in MATLAB (Mathworks, Inc.) function `procrustes`). We then visualized the average MDS results in a 2D space per ROI. In this space per ROI, for each category, 19 lines were drawn using the Procrustes transformed individual subject MDS position of that category. The line started in the dot of a category and ended at the coordinates (Procrustes transformed MDS results) of that category for the individual participant.

Results

Functional neuroanatomy of ventral OTC

Widespread and abundant activity to all categories of interest

We created a map of the functional neuroanatomy in response to the main conditions in our experiment. For each subject, the brain surface was reconstructed and results from several contrasts, calculated using the GLM, were projected onto this surface. In Figure 3 four exemplary subject surfaces are visualized (surfaces of all subjects can be found in Supplementary Figure 2). On each surface, selective activation (contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected) to several categories was displayed: to faces (orange), to bodies (ochre), to hands (yellow), to words (dark blue), to fake script (medium blue), to numbers (light blue), and to a typical objects category (chairs, medium green) to provide a reference landmark for our categories of interest.

Before zooming in on words specifically, we noted some general findings. First, in relation to category selectivity, we found category selectivity for each of these categories in all our subjects, thus for all types of characters and for all types of human-related categories (see Figure 3). This category selectivity was abundant and widespread. The selectivity for human-related and character categories appeared on the lateral side of the ventral surface, whereas the selectivity for the objects category appeared more medial, except in the posterior parts of the ventral surface (see Figure 3).

Second, we noted several findings in relation to overlaps between category selectivity. In Figure 3, overlap between selectivity for different human-related categories is depicted in red, for different character categories in purple, and for human-related categories overlapping with character categories in brown. These types of overlaps appeared all over the ventral surface in a predictable manner (see Figure 3): where selectivity for different human-related categories was nearby, we often also found clusters of overlap between different human-related category selectivity (shown in red), that seemed to act as a transition between the different category selectivity. This was similar for overlap between selectivity for character categories (shown in purple). This was also similar for the overlap (between selectivity for human-related and character categories (shown in brown): when selectivity for human-related and character categories appeared near each other, we often also found this type of overlap, suggesting a transition between the two kinds of selectivity. Overlap between selectivity for human-related categories overlapping with selectivity for the objects category is depicted in light green, and for character categories overlapping with the objects category in dark green. These kinds of overlap appeared in a small amount in the posterior parts of the ventral surface (see for example subject 14 in Supplementary Figure 2). The overlap between all types of selectivity, that is, selectivity for human-related, characters and the objects category, is depicted in pink. As expected, given that we defined

each category selectivity as one versus all other categories present in the experiment, this type of overlap rarely appeared and if it appeared, the area this covered on the surface was very small (see for example subject 1 in Figure 3).

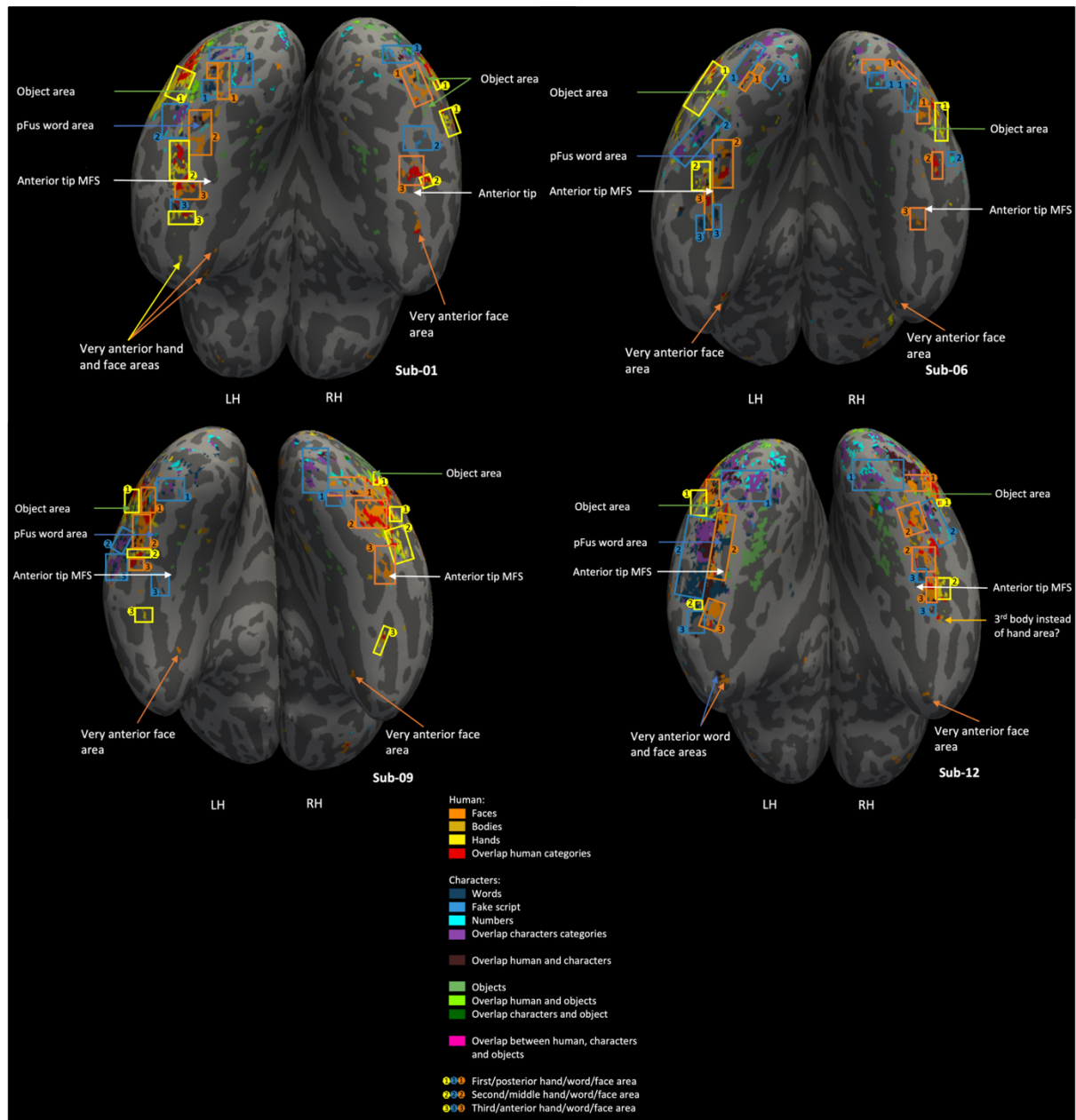


Figure 3. Category-selective regions shown upon annotated left and right hemisphere ventral surfaces of four example participants, together with an annotation of how to structure this category selectivity in three clusters of hand, word, and face selectivity. Annotations were made using a circle, each linked to a square, with the number 1, 2 or 3 to indicate the first, second or third hand (yellow), word (dark blue), and face (orange) area. They also include several arrows that point to the posterior object region (in green), the pFus word area (dark blue), the anterior tip of the mid-fusiform sulcus (white), a possible third body instead of the third hand area (ochre) and lastly, arrows pointing to even more anterior areas (color depends on the category). Color legend on the right. Selective activation (contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected) to faces in orange, to bodies in ochre, to hands in yellow, to words in dark blue, to fake script in medium blue, to numbers in light blue, and to an objects category (chairs) in medium green. Overlap between selectivity for different

human-related categories in red, overlap between selectivity for different character categories in purple, overlap between selectivity for human-related categories overlapping with character categories in brown, overlap between selectivity for human-related categories overlapping with selectivity for the objects category in light green, overlap between selectivity for character categories overlapping with the objects category in dark green, and overlap between selectivity for human-related, characters and the objects category in pink. LH: left hemisphere, RH: right hemisphere.

Several word-selective areas exist among areas selective to faces, hands and bodies in the left hemisphere

To understand where selectivity to word forms takes its place inside the map of functional neuroanatomy in response to our conditions, we first investigated where word selectivity (also called, the visual word form area (VWFA)) was located on the left hemisphere ventral surface and if we could consistently divide this activity into separate regions. Then, we investigated if this organization was consistent across subjects. Second, we located an important landmark in each subject (Weiner et al., 2014): the left hemisphere mid-fusiform sulcus (MFS). The anterior tip of this sulcus was proven to predict the location of a prominent area selective to faces: the FFA-2 or mFus, located on the middle fusiform gyrus (Weiner et al., 2014). By locating the MFS, we could investigate if this landmark also proved important for locating word selectivity and not just for face selectivity. Third, we wanted to explore if the organization of word selectivity and its possible subregions might be consistent in relation to the location of other category selectivity. We refer to Figure 3 for all anatomical descriptions described below.

First, we consistently found three clusters of word selectivity on the left hemisphere ventral surface of almost every subject (subject 4, 7, and 17 lacked the most anterior cluster and 17 also lacked the posterior cluster, Supplementary Figure 1 gives an overview of which areas were found in which subjects). The first cluster was located posterior, the second one more anterior, and the third even more anterior. The second cluster appeared in the middle of the first/posterior and third/anterior cluster but was often located more lateral than them. Thus, it seemed that word selectivity was present in at least three subregions that were consistently located across subjects. To quantify the replicability and size of the word selectivity, we performed a split-half analysis focused upon the middle and anterior ventral surface. The region of interest thus covered the second and third cluster of selectivity to words, faces, and hands. We were able to identify (in 15 subjects) word-selective voxels with half of the scans. Across these voxels, activity (as quantified by the beta value, see Methods section) to words in the other half of the scans was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to any of the other main categories of interest (see Figure 4A). The statistical details were as follows for the other categories of interest: words vs faces, $t(14)=11.66$, $p=1.35*10^{-8}$; words vs bodies, $t(14)=12$, $p=9.15*10^{-9}$; words vs hands $t(14)=12.63$, $p=4.84*10^{-9}$; and words vs objects (chairs) $t(14)=18.22$, $p=3.79*10^{-11}$. The word-selective voxels also showed a significantly higher response to words than to numbers ($t(14)=7.81$, $p=1.81*10^{-6}$) and to the fake script ($t(14)=11.64$, $p=1.38*10^{-8}$), even though the size of these differences was clearly smaller than when words were compared with objects categories. The division of this word selectivity into three subregions was helped by relating it to the selectivity for other categories. These categories are discussed in the following sections.

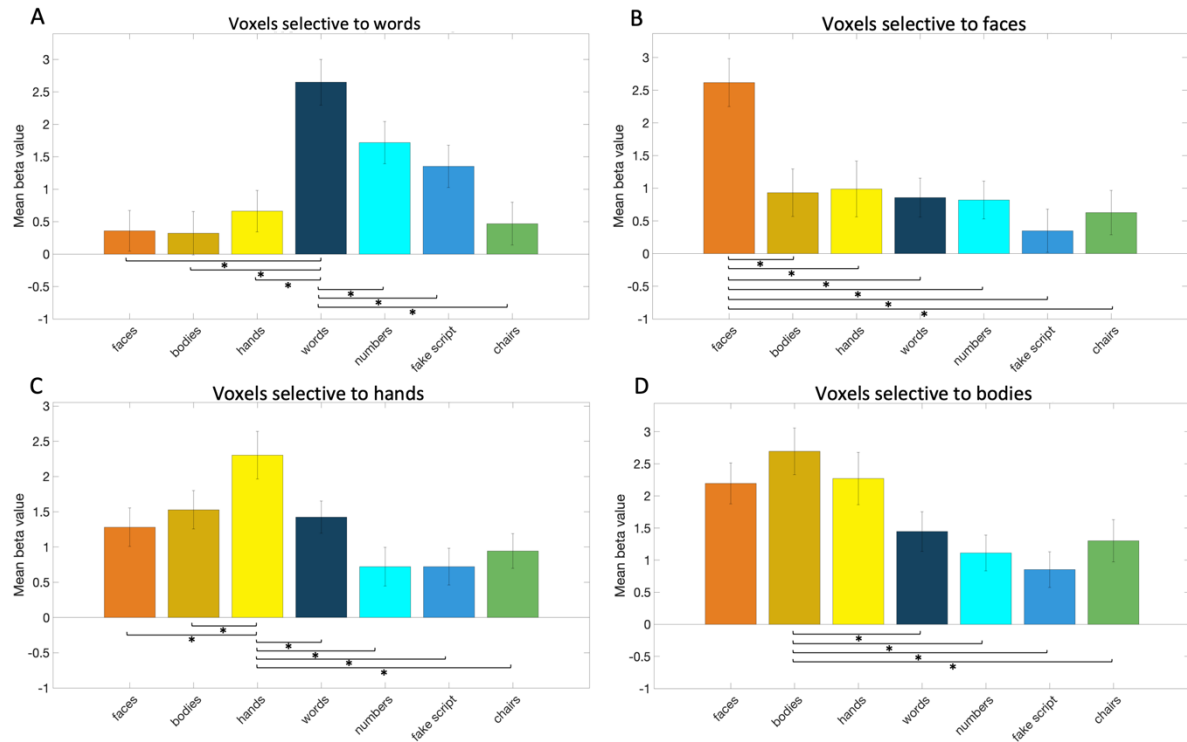


Figure 4. The response in word- (A), face- (B), hand- (C) and body- (D) selective voxels of the left hemisphere to seven categories (faces, bodies, hands, words, numbers, fake script, and objects: chairs). Responses were calculated using data that were independent from the data used to select the voxels (see Methods). Error bars represent the standard error, and the lines and stars indicate which paired t-tests between conditions were significant ($p < .008$).

Second, we then checked the relevance of the mid-fusiform sulcus (MFS) to locate the different word-selective clusters, building on top of prior demonstrations of its relevance to locate face selectivity. Using the location of the MFS, we determined where the three face areas in the left hemisphere were most likely located (as described in Weiner & Grill-Spector, 2010: on the inferior occipital gyrus (IOG); FFA1 or pFus, located on the posterior fusiform gyrus; and FFA2 or mFus, located on the mid-fusiform gyrus). For most subjects the selectivity for faces was plentiful and distributed over most of the ventral surface. Nonetheless, it was still possible, with guidance of the MFS, to indicate three clusters of face selectivity, suited to be defined as IOG, pFus and mFus, in almost all subjects (subject 13's anterior cluster seemed too anterior to be mFus, subject 14 lacked the pFus, subject 15 and subject 17 lacked the IOG). In some subjects the separation between pFus and mFus was not straightforward (subject 16 and subject 19). The replicability of face selectivity overall was tested with the split-half analysis. We were able to identify (in 14 subjects) face-selective voxels across the middle and anterior ventral surface with one half of the scans. Across these voxels, activity (as quantified by the beta value, see Methods section) to faces in the other half of the scans was

significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to any of the other main categories of interest (see Figure 4B). The statistical details were as follows: faces vs bodies, $t(13)=12$, $p=2.07*10^{-8}$; faces vs hands, $t(13)=6.77$, $p=1.32*10^{-5}$; faces vs words, $t(13)=6.47$, $p=2.11*10^{-5}$; faces vs numbers, $t(13)=9.16$, $p=4.91*10^{-7}$; faces vs fake script, $t(13)=9.98$, $p=1.85*10^{-7}$; and faces vs objects (chairs), $t(13)=8.98$, $p=6.19*10^{-7}$. The MFS did not make a critical contribution to the localization of the three subregions of the VWFA, although a few relationships were often present. We noticed that the middle subregion of the VWFA was lateral and more posterior in comparison to the anterior tip of the MFS. Additionally, we found that the anterior subregion of the VWFA appeared in the neighborhood of the mFus, which was often in line with the location of the anterior tip of the MFS as previously mentioned.

Third and last, we looked at the selectivity to the other categories besides words and faces. We found that there were at least two, sometimes three (in subject 1, 9, 10, 13, and 19), clusters of selectivity to hands. On the surfaces of some subjects (subject 3, 7, 8 and 18) we could not determine the third anterior hand cluster. However, these subjects did instead show an anterior cluster of selectivity to bodies in a location where we expected the third hand cluster to reside. The replicability of hand selectivity overall was tested with the split-half analysis. We were able to identify (in 16 subjects) hand-selective voxels across the middle and anterior ventral surface. Across these voxels, activity (as quantified by the beta value, see Methods section) to hands was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to any of the other main categories of interest (see Figure 4C). The statistical details were as follows: hands vs words, $t(15)=3.64$, $p=.002$; hands vs numbers, $t(15)=8.19$, $p=6.47*10^{-7}$; hands vs fake script, $t(15)=9.58$, $p=8.78*10^{-8}$; and hands vs objects (chairs), $t(15)=6.3$, $p=1.43*10^{-5}$. The hand-selective voxels also showed a significantly higher response than to faces ($t(15)=3.3$, $p=.005$) and to bodies ($t(15)=3.85$, $p=.002$), but the size of these differences was clearly smaller than when hands were compared with character categories, with the exception of the words category, or objects categories. The hand-selective regions were the most helpful in delineating the three word-selective regions, which will be explained in Section 3.1.2.1.

Given that body selectivity has often been used to obtain a detailed picture of face selectivity (e.g., Weiner & Grill-Spector, 2010), we looked at the body-selective areas on the ventral surface and found them to lie separately from the hand areas. We also identified body-selective voxels across the middle and anterior ventral surface using the split-half analysis. Across these body-selective voxels (identified in 15 subjects), activity (as quantified by the beta value, see Methods section) to bodies was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to words ($t(14)=6.31$, $p=1.93*10^{-5}$), numbers ($t(14)=5.84$, $p=4.29*10^{-5}$), fake script ($t(14)=7.51$, $p=2.84*10^{-6}$) and objects (chairs) ($t(14)=7.41$, $p=3.32*10^{-6}$) (see Figure 4D). However, activity to bodies was not significantly

different from faces (significant only at uncorrected threshold: $p < .05$; $t(14)=2.89$, $p=.01$) or hands ($t(14)=1.51$, $p=.15$). On the contrary, in the hand-selective voxels, the activity to hands was significantly higher than to faces or bodies (described above). To compare the activity to its preferred category and second-preferred category, we compared the hand- and body-selective voxels/areas in a paired t-test. Specifically, we compared the difference between bodies and hands in the body area with the difference between hands and bodies in the hand area. We did not find a significant difference, although there was a trend ($t(12)=-1.45$, $p=.17$) that the difference between bodies and hands in the body area tended to be smaller than the difference between hands and bodies in the hand area. This is all the more reason to consider hand selectivity as an important property also in ventral occipitotemporal cortex.

With the split-half analysis, we could also identify voxels selective to numbers and to fake script across the middle and anterior ventral surface, but there was no evidence for more selectivity to numbers than words or to fake script than words, in the number and fake script areas. First, across these number-selective voxels (identified in 17 subjects), activity (as quantified by the beta value, see Methods section) to numbers was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to faces ($t(16)=8.27$, $p=3.57*10^{-7}$), bodies ($t(16)=10.15$, $p=2.23*10^{-8}$), hands ($t(16)=10.34$, $p=1.72*10^{-8}$), fake script ($t(16)=3.16$, $p=.006$), and objects (chairs) ($t(16)=9.01$, $p=1.15*10^{-7}$) (see Supplementary Figure 3A). It was not significantly different from words ($t(16)=-1.22$, $p=.24$). Second, across these fake script-selective voxels (identified in 17 subjects), activity (as quantified by the beta value, see Methods section) to fake script was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to faces ($t(16)=4.77$, $p=.0002$), bodies ($t(16)=5.27$, $p=7.60*10^{-5}$), hands ($t(16)=4.42$, $p=.0004$), and objects (chairs) ($t(16)=3.57$, $p=.003$) (see Supplementary Figure 3B). It was significantly lower than words ($t(16)=-3.51$, $p=.003$) and not different from numbers ($t(16)=-1.37$, $p=.19$). Lastly, given that on the map of the ventral surface, we used selectivity to a typical objects category (chairs) as a reference landmark, we also attempted to identify object-selective voxels using the split-half analysis. We identified these object-selective voxels in all 19 subjects and found that activity (as quantified by the beta value, see Methods section) to objects was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to words ($t(18)=3.38$, $p=.003$), numbers ($t(18)=4.38$, $p=.0004$) and fake script ($t(18)=4.39$, $p=.0004$), but not different (significant only at uncorrected threshold: $p < .05$) from faces ($t(18)=2.70$, $p=.01$), bodies ($t(18)=2.75$, $p=.01$) and hands ($t(18)=2.8$, $p=.01$) (see Supplementary Figure 3C).

The location of hand selectivity serves as a reference point to locate word and face selectivity in the left hemisphere

The clusters of selectivity to hands were so consistent in location that they were an important reference point for locating the subregions of the VWFA and the typical face areas (IOG, pFus and mFus) in most of the subjects. First, for the posterior part of the ventral surface, we found that the first/posterior hand region, which lied lateral and posterior on the ventral surface, was an important reference landmark. This hand-selective region was previously described by Bracci and colleagues (2010) as a region separate from the extrastriate body area (EBA) and selective to hands more than to whole bodies and other body parts. Lateral to this first hand region and more medial on the surface, we found a face-selective cluster (IOG) and a first word-selective cluster of activity, around this IOG region. In many of the subjects, on top of this word cluster around IOG, there was selectivity to words and other character categories distributed across the posterior ventral surface, also lateral to this first hand region. In some of the subjects, we also found that a cluster of selectivity to objects and/or a cluster of overlap of object selectivity with character and/or human-related selectivity, separated this first hand cluster from the first face and word cluster (the most clearly visible in subject 1, 3, 6, 9, 11, 12 and 18).

Second, for the middle part of the ventral surface, the second/middle cluster of selectivity to hands, which lied more anterior on the ventral surface than the first hand cluster, also proved to be a useful reference landmark. In between the first and second cluster of hand selectivity, the second word-selective cluster of activity was located, with more medial to that word cluster, the second face selectivity cluster (pFus). Due to the central location of the second selectivity cluster to hands, we decided to examine this selectivity also in the volume space (Figure 5). We found that all selectivity to categories on and around this location of the second hand cluster lined up with the anatomical structure of the brain. We also found that pFus seemed to be joined, or sometimes broken up into two, by another cluster of selectivity to words and/or other characters (visible in subject 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 15, and 16). We called this the pFus word area. To investigate this more closely, we also looked at this area in the volume space (Figure 6) in two of our subjects. In subject 1, the pFus word area seemed to break the pFus area in two, whereas in subject 11, the pFus word area adjoined the pFus area in the sulcus.

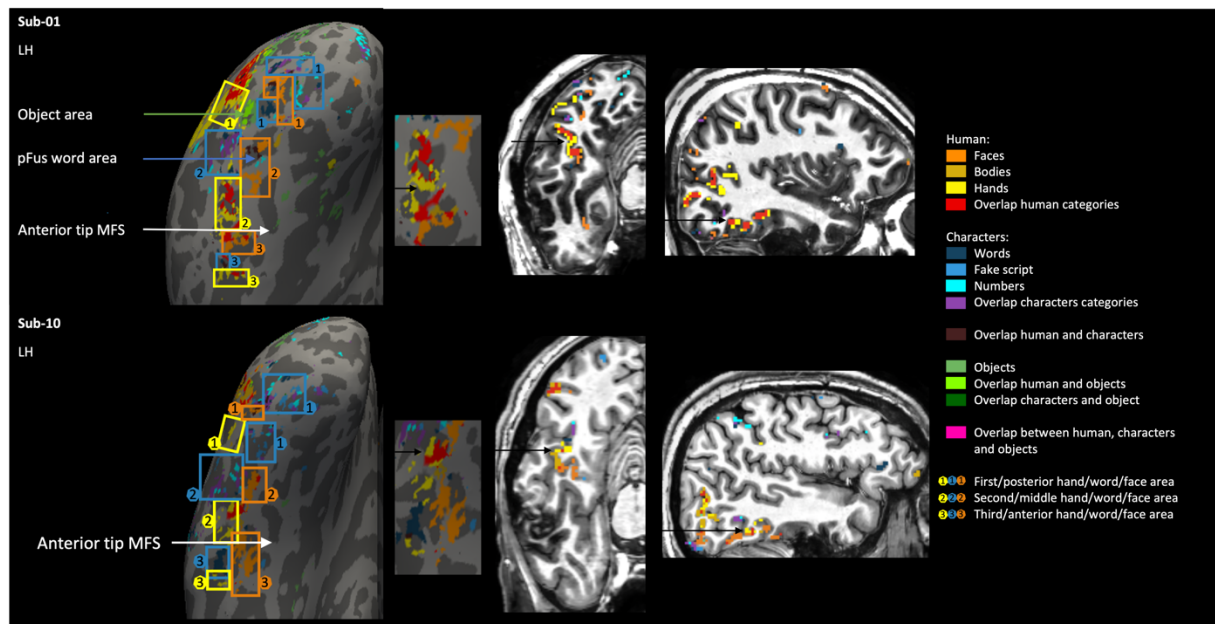


Figure 5. Visualization of the second/middle hand area of subject 1 (top) and subject 10 (bottom) in surface space on the left, zoomed in visual on the surface in the middle, and then on the right in volume space in axial and sagittal view. Black arrows on the zoomed in surface in the middle indicate the location used to visualize the volume space on the right (this location is also indicated by black arrows on the volume on the right).

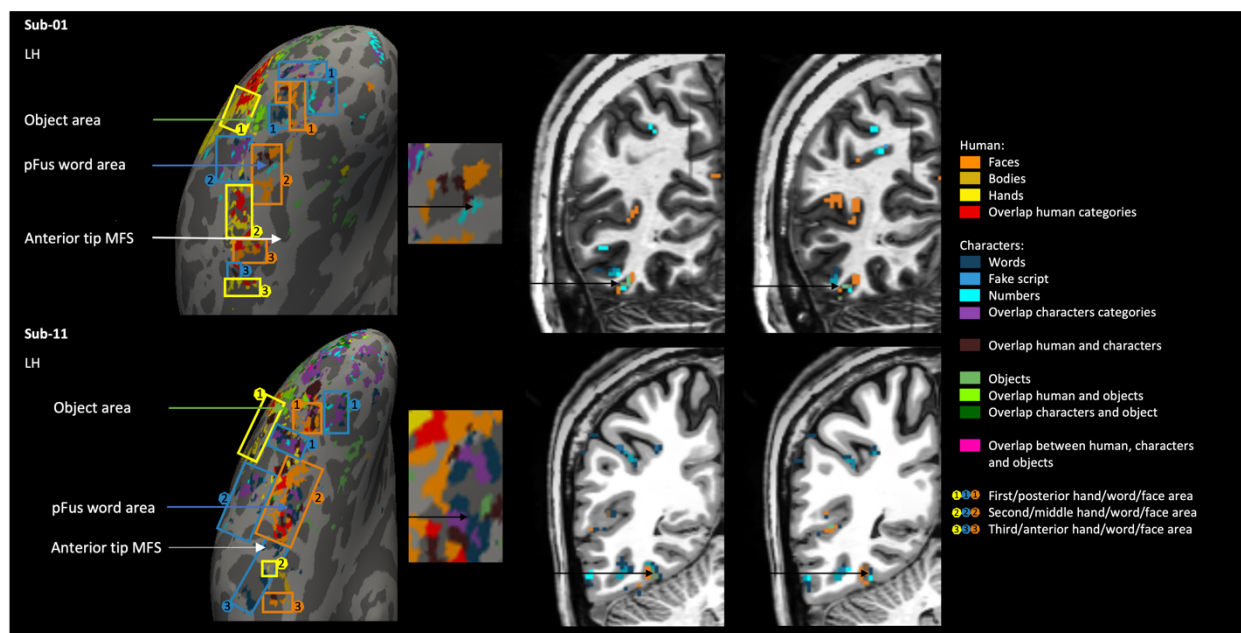


Figure 6. Visualization of the pFus word area of subject 1 (top) and subject 11 (bottom) in volume space in surface space on the left, zoomed in visual on the surface in the middle, and then on the right in volume space in two adjoining slices in coronal view. Black arrows on the zoomed in surface in the middle indicate the location used to visualize the volume space on the right (this location is also indicated by black arrows on the volume on the right).

Third, we looked at the anterior part of the ventral surface. Like mentioned above, in a couple of the subjects, there seemed to be a small third/anterior cluster of selectivity to hands and/or bodies. Around the second hand cluster, in between the second and the third hand selectivity cluster (when present), we found the third face (mFus) and word selectivity cluster. Looking even more anterior, we also regularly found very anterior face selectivity clusters: in subject 1 (accompanied by a hand selectivity cluster), 3, 4 (+ hand cluster), 5 (accompanied by a word selectivity cluster), 6, 9, 10, 12 (+ word cluster), 13 (if this is not in fact mFus, see above, + word cluster), 16, and 19.

Less word selectivity in the right hemisphere

It has been shown before that the left hemisphere showed the most prominent selectivity to words and thus studies have often analyzed only the left hemisphere. After investigating the left hemisphere of our subjects, we sought out to compare it to the selectivity found in the right hemisphere. Specifically, we compared if the standard organization of the left hemisphere selectivity clusters was similar in the right hemisphere. This consisted of first, checking if all the same types of selectivity were present and if the same number of clusters of each type of selectivity appeared. Second, we examined if these areas were organized in a similar way. We refer to Figure 3 for all anatomical descriptions described below. Overall, we found the same types of selectivity in the right hemisphere as in the left hemisphere. Some types of selectivity were often less clear in appearance in the right versus the left hemisphere (e.g., smaller, more distributed, or a somewhat strange or inconsistent location in many subjects).

We were able to consistently find selectivity to words in the posterior region and often also in the middle region (except 9, 10, 11, 16, 18 and 19) of the ventral surface. In several subjects, this selectivity to characters did not include selectivity to words specifically, and this was true for all word regions. Instead, the selectivity was mostly generalized to numbers and/or fake script (posterior region: see subject 1, 5, 8, 14, 16 and 18; middle region: see subject 1, 2, 6, 8, and 14). This was also evidenced by the split-half analysis, described below, that showed no difference between selectivity to words, numbers, or fake script within the word-selective voxels of the middle and anterior ventral surface. The third cluster of selectivity to words could only be found in some subjects (subject 5, 12 and 18). In subject 2, 8 and 9 there was a small selectivity cluster to the numbers character category and in subject 2, there was also a lateral (in left hemisphere this appeared medial) cluster of selectivity to numbers. The replicability of word selectivity overall was tested with the split-half analysis and we were able to identify (in 11 subjects) word-selective voxels across the middle and anterior ventral surface (see Figure 7A). Across these voxels, activity (as quantified by the beta value, see Methods section) to words was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to faces ($t(10)=4.65$,

$p=.0009$), and bodies ($t(10)=5.06$ $p=.0005$), but not different from (significant only at uncorrected threshold $p < .05$) hands ($t(10)=2.88$, $p=.02$), fake script ($t(10)=2.83$, $p=.02$) and objects/chairs ($t(10)=3.18$, $p=.01$). It was also not significantly different from numbers ($t(10)=2.05$, $p=.07$). To further look into the lack of difference in selectivity between words, numbers and fake script in these right-hemisphere word-selective voxels, as opposed to the left-hemisphere word-selective voxels, we performed two paired t-tests to directly compare the two hemispheres. We applied Bonferroni correction at $p < 0.03$. In the first test we compared the difference between words and numbers between the left and right hemisphere word-selective voxels, in the second we compared the difference between words and fake script. Of note to these comparisons is that in the left hemisphere, we could not identify word-selective voxels in 4 of the 19 subjects, whereas in the right hemisphere we could not identify such voxels in 8 of the 19 subjects. We found that in the left hemisphere, the difference between words and numbers ($t(7)=2.6$, $p=.04$), and between words and fake script ($t(7)=3.78$, $p=.007$), was significantly bigger (marginally insignificant in case of words versus numbers) as compared to the right hemisphere word-selective voxels.

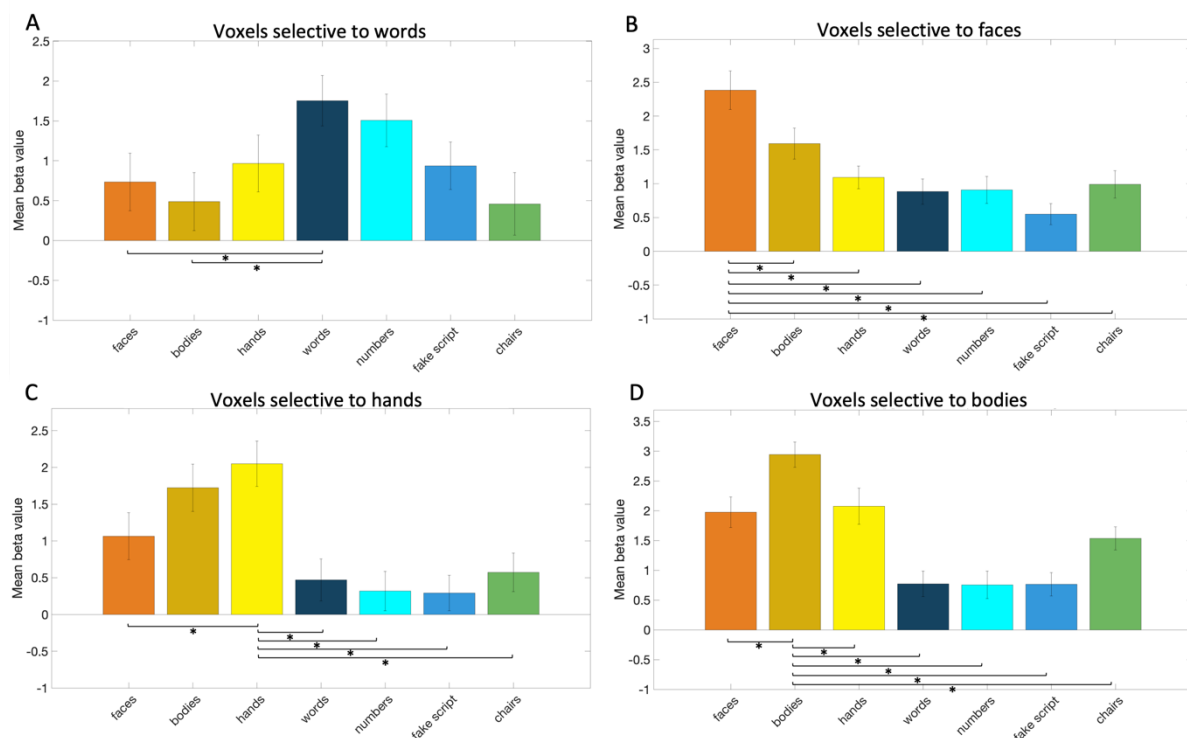


Figure 7. The response in word- (A), face- (B), hand- (C) and body- (D) selective voxels of the right hemisphere to seven categories (faces, bodies, hands, words, numbers, fake script, and objects: chairs). Responses were calculated using data that were independent from the data used to select the voxels (see Methods). Error bars represent the standard error, and the lines and stars indicate which paired t-tests between conditions were significant ($p < .008$).

Like in the left hemisphere, we also located the anterior tip of the mid-fusiform sulcus in the right hemisphere and again found, that it did not make a critical contribution to the localization of the three subregions of the VWFA. With guidance of the MFS, we could indicate three clusters of face selectivity, suited to be defined as IOG, pFus and mFus in most subjects. Regarding the IOG, all but one subject (subject 10) showed the IOG cluster of face selectivity. Some subjects (1, 3, 4, 5, 10, 14) did not show a pFus area at all or the activity seemed more likely to still be a part of the first face selectivity cluster (IOG) and/or the third face selectivity cluster (mFus). Like in the left hemisphere, the separation from the third/mFus area was not always clear (e.g., subject 7, 8 and 18). The third cluster of selectivity to faces (mFus) was detected in most subjects (except subject 10, 11 and 17). With the split-half analysis, we were able to identify face-selective (in 11 subjects) voxels across the middle and anterior ventral surface. Across these voxels, activity (as quantified by the beta value, see Methods section) to faces was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to bodies ($t(10)=6.55$, $p=6.46*10^{-5}$), hands ($t(10)=6.74$, $p=5.08*10^{-5}$), words ($t(10)=8.45$, $p=7.24*10^{-6}$), numbers ($t(10)=11.24$, $p=5.38*10^{-7}$), fake script ($t(10)=8.61$, $p=6.13*10^{-6}$), and objects (chairs) ($t(10)=9.76$, $p=1.99*10^{-6}$).

Like in the left hemisphere, we often found two, but rarely three, areas of selectivity to hands. The posterior hand area appeared in all the subjects in the right hemisphere. The second cluster of selectivity to hands was apparent in most of the subjects (except subject 4, 5, 6, 7, 10 and 11). The third cluster of selectivity to hands could not be determined in most of the subjects (only in subject 9, 11, and maybe 17). Also alike the left hemisphere, we could instead around this location find selectivity to bodies in some subjects (subject 3, 12, 14 and 16). With the split-half analysis, we were able to identify (in 16 subjects) hand-selective voxels across the middle and anterior ventral surface. Across these voxels, activity (as quantified by the beta value, see Methods section) to hands was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to faces ($t(15)=4.5$, $p=.0004$), words ($t(15)=7.52$, $p=1.83*10^{-6}$), numbers ($t(15)=7.20$, $p=3.06*10^{-6}$), fake script ($t(15)=8.52$, $p=3.92*10^{-7}$), and objects (chairs) ($t(15)=8.96$, $p=2.07*10^{-7}$). It was not significantly different from bodies (significant only at uncorrected threshold: $p < .05$, $t(15)=2.35$, $p=.03$) (see Figure 7C). A big difference between the left and right hand areas was the strength of the selectivity for words. Indeed, the difference between hands and words was smaller in the left hand-selective voxels as compared to the right hand-selective voxels ($t(14)=-2.78$, $p=.02$), suggesting that the left hand area is more selective to words than the right hand area.

Regarding the body areas, we again found those separate from the hand areas on the ventral surface. We identified (in 14 subjects) body-selective voxels across the middle and anterior ventral surface using the split-half analysis. We found that activity (as quantified by the beta value, see Methods section) to bodies was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than

to faces ($t(13)=6.73$, $p=1.42*10^{-5}$), to hands ($t(13)=4.31$, $p=.0009$), in contrast to the left hemisphere body-selective voxels. The activity to bodies was also significantly higher than to words ($t(13)=10.19$, $p=1.44*10^{-7}$), numbers ($t(13)=9.28$, $p=4.27*10^{-7}$), fake script ($t(13)=10.5$, $p=1.01*10^{-7}$) and objects (chairs) ($t(13)=10.13$, $p=1.55*10^{-7}$) (see Figure 7D).

With the split-half analysis, we could also identify voxels selective to numbers and to fake script across the middle and anterior ventral surface, but there was no evidence for more selectivity to numbers than words or to fake script than words, in the number and fake script areas. First, across these number-selective voxels (identified in 14 subjects), activity (as quantified by the beta value, see Methods section) to numbers was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to faces ($t(13)=5.09$, $p=.0002$), bodies ($t(13)=4.73$, $p=.0004$), hands ($t(13)=4.18$, $p=.001$), fake script ($t(13)=4.31$, $p=.0008$), and objects ($t(13)=4.29$, $p=.0009$) (see Supplementary Figure 4A). It was not significantly different from words ($t(13)=-0.37$, $p=.72$). Second, across these fake script-selective voxels (identified in 11 subjects), activity (as quantified by the beta value, see Methods section) to fake script was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to faces ($t(10)=10.78$, $p=7.97*10^{-7}$), bodies ($t(10)=4.83$, $p=.0007$), hands ($t(10)=5.61$, $p=.0002$), and objects ($t(10)=3.41$, $p=.007$) (see Supplementary Figure 4B). It did not differ significantly from words ($t(10)=-1.73$, $p=.11$) and from numbers ($t(10)=-0.56$, $p=.59$). Lastly, given that on the map of the ventral surface, we used selectivity to a typical objects category (chairs) as a reference landmark, we also attempted to identify object-selective voxels using the split-half analysis. Across these object-selective voxels (identified in 15 subjects), activity (as quantified by the beta value, see Methods section) to objects was significantly (i.e., $p < .008$ based on Bonferroni correction) higher than to faces ($t(14)=3.13$, $p=.007$) and to numbers ($t(14)=4.11$, $p=.001$), but did not differ significantly (significant only at uncorrected threshold: $p < .05$) from hands ($t(14)=2.29$, $p=.04$), words ($t(14)=2.69$, $p=.02$) and fake script ($t(14)=2.66$, $p=.02$). It was also not different from bodies ($t(14)=0.95$, $p=.36$) (see Supplementary Figure 4C).

The location of hand selectivity cannot serve as a reference point to locate word and face selectivity in the right hemisphere

Then, we investigated if the hand areas in this hemisphere also allowed to consistently locate the VWFA subareas and the typical face areas, like in the left hemisphere. We did this by first comparing the locations of these areas in the right hemisphere to the locations of the left hemisphere, and then looked at the possible guidance of the hand areas in locating the other areas. In conclusion, we found that the hand areas could not so much serve as a reference landmark like they did in the left hemisphere, in part due to the more complex nature of the hand and word selectivity in the right

hemisphere as compared to the left hemisphere. For more details, we refer to the supplementary results.

The organization of word- and other category-selective areas in left-handed subjects

Three subjects in our subject pool were left-handed: subject 2, 7 and 17. We refer to Supplementary Figure 2 for all results described below. When looking at the left hemisphere, we found that subject 2 and subject 7 showed the same organization in their left hemisphere as the right-handed subjects did, as described above. However, subject 17 differed from this. In the left hemisphere, we found the first hand cluster, the second word cluster anterior to that, the pFus cluster medial to that and the second hand cluster below it. There also seemed to be a mFus area present in this subject 17. All the other clusters were missing: the first (unless very small and distributed) and third word cluster, the IOG cluster, and the third hand cluster.

When looking at the right hemisphere of the left-handed subjects, in subject 17, we could find an organization of areas somewhat like the one we found in the left hemisphere of our right-handed subjects. The subject showed all first and second hand, word, and face areas. In the right hemisphere of subject 2, we found the standard organization of selectivity like that of the left hemisphere in the right-handed subjects. We found the first hand cluster, the first word cluster, the IOG cluster, the second hand cluster, the second word cluster, pFus, and mFus. This means that we did not clearly see a third cluster of selectivity to words. The right hemisphere organization of selectivity of subject 7 seemed similar but less clear than that of subject 2.

Character categories group together in representational space and words do not show a clear competition with other relevant categories

To achieve a high-resolution mapping of word selectivity relative other visual categories, we investigated the data in two ways: the functional neuroanatomy and the representational space. For every subject, two different matrices were created through applying MVPA: for the occipitotemporal cortex (OTC) ROI, separately for the left and the right hemisphere. Each of these matrices were normalized and then averaged across all subjects. These matrices were reliable: every subject's matrix was significantly correlated with the average matrix (mean correlation for the left OTC ROI: $r = .80$, mean correlation for the right OTC ROI: $r = .74$).

These matrices were then visualized in a two-dimensional space using MDS. We called this result the representational space of our categories (see Figure 8). Within the visualization of the matrices in a 2D space, categories that lay close together shared a more similar neural representation than categories that lay further away from each other. Inside this space, we also visualized the Procrustes

transformed MDS results of every participant for each category using one line per participant and found that overall, our subjects presented with consistent results. We looked first at the results from the left OTC ROI, without focusing specifically on words. We noted some general findings: a grouping of character categories (words, numbers, fake script), a grouping of animate categories (faces, bodies, cats, hands), and a group of inanimate categories in the middle of the plot with the following categories: cars, fish, chairs, instruments, flowers, vegetables, hammers, scissors, buildings, trees, cubies. Fish are animate shapes but in MVPA results they are close together with inanimate categories (Connolly et al., 2012). Cubies and smoothies were unknown shapes for the participants. The cubies fell with the inanimate group, whereas the smoothies could be included within the animate group in the plot. This might be explained by differences in mid-level features between the two categories, based on previous research results: animate categories have a higher curvilinearity than artifacts (Levin et al., 2001), perceived curvature predicts the classification of animacy of a texform and the curvature differences between animals and artifacts can explain to some degree the mixed-animacy search advantage (Long et al., 2017). We found that the control condition (scrambled) lay separate from all other categories. When looking at the right OTC ROI, we found similar results.

Then, we focused on understanding the position of words and other character categories in particular, inside this representational space. Characters grouped together. Using paired t-tests (i.e., $p < .006$ based on Bonferroni correction), we investigated if numbers and fake script indeed were significantly more similar to words than other seemingly close categories like cars and faces. The distance (based on the normalized LDC value of every subject) between words and numbers was significantly smaller than the distance between words and cars in both the left ($t(18)=-10.81$, $p=2.64*10^{-9}$) and the right hemisphere ($t(18)=-5.84$, $p=3.32*10^{-5}$). The same was true when comparing the distance between words and numbers with the distance between words and faces in both the left ($t(18)=-11.06$, $p=1.86*10^{-9}$) and the right hemisphere ($t(18)=-12.92$, $p=1.53*10^{-10}$). To conclude, numbers were significantly more similar in its neural activation pattern to that of words, as compared to other categories that seemed the closest in the MDS space. We then investigated the same for the other character category: fake script. The distance between words and fake script was significantly smaller than the distance between words and cars in both the left ($t(18)=-7.07$, $p=1.37*10^{-6}$) and the right hemisphere ($t(18)=-4.18$, $p=.0006$). The same was true when comparing the distance between words and fake script with the distance between words and faces in both the left ($t(18)=-10.09$, $p=7.82*10^{-9}$) and the right hemisphere ($t(18)=-12.26$, $p=3.55*10^{-10}$). As a conclusion, both numbers and fake script lied significantly closer to words than other categories.

Second, we focused, as explained in the Introduction, on the position of human-related categories of interest in relation to the position of words. We performed paired t-tests (i.e., $p < .01$ based on

Bonferroni correction) to investigate if the neural activation pattern of faces was more similar to that of words than bodies and hands, based on the idea of their competition and the observation that bodies and hands lay further away in Figure 8. The distance between words and faces was not significantly smaller than the distance between words and bodies in the left ($t(18)=-1.18$, $p=.25$) hemisphere, but it was significantly smaller in the right hemisphere ($t(18)=-4.31$, $p=.0004$). When comparing the distance between words and faces to the distance between words and hands, we found that the distance between words and faces was not significantly larger than the distance between words and hands in the left hemisphere ($t(18)=-0.38$, $p=.71$), nor in the right hemisphere ($t(18)=1.63$, $p=.12$). To summarize, in the left hemisphere, faces was not the closest category to words: bodies/hands were as close to words as faces was. In the right hemisphere, faces was also not the closest category to words: faces was closer to words than the bodies, but hands was as close to words as faces was. This suggests, if using distance between categories as a measure for their competition, that the competition between words and faces may not be specific to faces: the competition may extend to other human-related categories.

Third, we sought out the closest inanimate category (cars) in the space and compared its distance to words with the distance of the closest animate category (faces) to words. The distance between words and faces was not significantly different from the distance between words and cars (left hemisphere: $t(18)=3.94$, $p=.22$; right hemisphere: $t(18)=6.07$, $p=.15$). This suggests, if using distance between categories as a measure for their competition, that the faces category did not compete more with the words category than the cars category.

Finally, we compared the left to the right hemisphere. When looking at the functional neuroanatomy of the ventral OTC, we found that the right hemisphere, as compared to the left hemisphere, showed a weaker selectivity to words. Word-selective areas, identified using the split-half analysis, were not significantly more responsive to words than to other types of characters in the right hemisphere. In addition, in most of the subjects we could not localize the third/anterior cluster of word selectivity on the right ventral brain surface. On the other hand, the organization in the representational space of the left and right hemisphere (Figure 8A and 8B) seems very similar. We sought out to understand if the left and right hemisphere processed words and other characters in a similar way in terms of representational similarity. To this end, we compared the normalized LDC value of each subject for several pairs of categories between the left and right hemisphere: for words versus numbers and for words versus fake script. We found that one of the two tests was almost significant (i.e., $p < .03$ based on Bonferroni correction): words versus numbers ($t(18)=2.25$, $p=.04$). The other test, words versus fake script, was significant ($t(18)=2.53$, $p=.02$). Thus, the neural pattern of words differed more from the pattern of numbers and fake script in the left hemisphere than in the right

hemisphere. This was in accordance with the results from the split-half analysis, where the left hemisphere differentiated between words and other characters, but the right hemisphere did not.

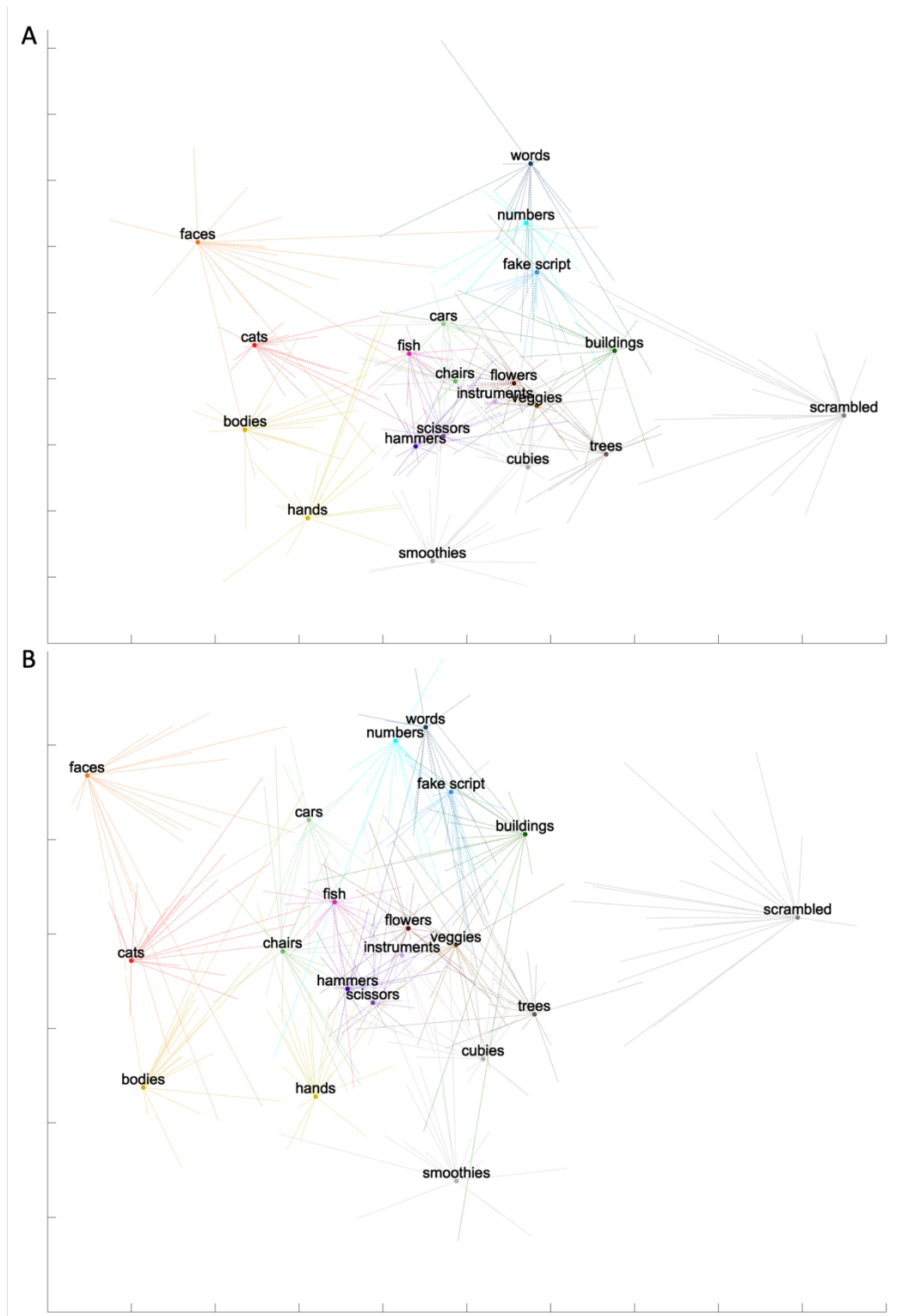


Figure 8. Visualization of the average MVPA matrix of the (A) left OTC ROI and (B) right OTC ROI, using MDS. MDS was also applied on each individual MVPA matrix and these MDS results were then Procrustes transformed to the average MDS result. The individuals' Procrustes transformed MDS results are depicted by one line per subject for every category within the space.

Discussion

In this study, we investigated with a high spatial resolution where visual word forms (in the format of letter strings) are located on the ventral surface map of the functional neuroanatomy of the occipitotemporal cortex (OTC). We related this to other category-selective areas predicted to be relevant based on the neural recycling theories and based on the proposal by Yeatman and colleagues (2021, see Introduction) (other character categories: numbers and fake script, and human-related categories: faces, hands, and bodies). In addition, we also investigated word selectivity in the representational space and if a competition between words and faces (and in extension possibly also hands and bodies) was also apparent here. We scanned 19 participants with 7T fMRI and presented them shapes of 20 different categories. We performed our analyses at the level of the individual brain (thus, without spatial normalizing or smoothing).

We mapped the functional neuroanatomy of OTC, by visualizing the selective activation to characters and human-related categories on the ventral surface of the brain and by showing its replicability and the strength of category selectivity through a split-half analysis. We also chose one typical objects category (chairs) as a reference. We found abundant and wide-spread selectivity to the categories of interest across the ventral surface.

In the left hemisphere of ventral OTC, three clusters of word selectivity were located in a consistent manner across participants. One was found posterior and covered an area lateral to medial, one was located more in the middle of the ventral surface and mostly lateral, and one lied anterior and also rather lateral. In addition, we found an extra cluster of word selectivity that we called the pFus word area because it lied either in between or next to the pFus area of face selectivity. Comparing these word areas to the functional subdivisions made in recent research, we believe the middle/second and the anterior/third word area to correspond with the pOTS/VWFA-1 and mOTS/VWFA-2 respectively (Lerma-Usabiaga et al., 2018; White et al., 2019). The posterior/first word area corresponds to posterior regions that have been sometimes defined separately from the VWFA, based on the functional observation that these regions are selective for letters but do not specifically prefer letter combinations that form words (James et al., 2005; Vinckier et al., 2007; Wong et al., 2009; Strother et al., 2015; Yeatman et al., 2021). As far as we know, the pFus word area has not before been defined as a subarea of the VWFA. Since previous research has predominantly subdivided the VWFA based on functional differences, it could be that the pFus word area could not be discriminated from pOTS/VWFA-1 and/or because these studies did not localize the separate face areas. A recent study by Boring and colleagues (2021) localized word and face selectivity at high spatial resolution in each individual subject. When comparing our results to theirs, we believe that on the brain surface of most of their subjects, the pFus word area could be defined, as could the other three word areas (the

posterior, middle and anterior word area). Word selectivity within the left hemisphere was replicated in the split-half analysis, which also showed that the selectivity to words was strong in word-selective voxels: significantly stronger than to all other categories of interest and the strongest when comparing words to non-character categories. Thus, as expected, the word-selective areas also showed some selectivity to other character related categories (numbers and fake script), but significantly less than to words.

Given the relevance of the mid-fusiform sulcus to the location of face-selective areas (Weiner et al., 2014), we located this sulcus in all our subjects and investigated if it was also relevant to the word-selective areas that we found in the left hemisphere of OTC. It did not make a critical contribution to the locating of these word areas, but it could help in locating the several face areas (IOG, pFus and mFus). We replicated the face selectivity with the split-half analysis and showed that the selectivity to faces in the face areas was stronger than to any of the other categories of interest.

Unexpectedly, in the left hemisphere, we also found at least two clusters of selectivity to hands, one lateral and posterior, which has been demonstrated before by Bracci and colleagues (2010), and one more in the middle of the ventral surface of OTC. This hand selectivity was also replicated in the split-half analysis, which also showed that in these hand areas, selectivity to hands was stronger than to any of the other categories of interest. Interestingly, the difference in activity between hands and words, faces, and bodies, was smaller than between hands and objects, numbers, and fake script. This is in line with the recent results of Nordt and colleagues (2021) that suggested a competition between limb and word selectivity: we found that areas selective to hands also showed some selectivity to words. Apart from these hand-selective areas, we also found body-selective areas on the ventral surface of the left hemisphere. The split-half analysis proved its replicability and showed that these body-selective areas also showed strong selectivity to hands and to faces (selectivity to bodies was not significantly stronger than to hands or faces). An extra test suggested a trend that the hand-selective areas showed a more distinct selectivity to its preferred category (hands) than the body-selective areas did to bodies.

More surprising, these hand-selective areas provided a reference point from which we could easily locate the word- and face-selective areas on the left hemisphere ventral surface of OTC. Medial to the first hand area, lied the first word and face area. The second hand area lied more anterior than the first one around the middle of the surface. In between this first and second hand area, the second word area was located and more medial to that, the pFus face area, which was joined or broken up by what we called the pFus word area. In some of the subjects there was also a third hand area even more anterior on the ventral surface of OTC. Below/around the second hand area and in between this second and the third hand area, we found the mFus face area and the third word area. These observations are

in accordance with the proposal by Yeatman and colleagues (2021) (which was in part based on Grill-Spector & Weiner, 2014): that the VWFA-1 (the middle word area in our data) and VWFA-2 (the anterior word area in our data) would lie next to the pFus and mFUs face area, and that a body/limb area would subdivide between the VWFA-1 and VWFA-2. We found that the hand areas specifically, more than the body areas, could guide the subdivisions of word (and face) areas. Additionally, we observed not one hand area (in between VWFA-1 and VWFA-2) but two (or even three) hand areas, that could also guide the subdivision between the first/posterior word area and the VWFA-1 (second/middle word area).

Regarding the right hemisphere ventral surface of OTC, we could identify the first and often also the second word-selective cluster. We could not identify the third/anterior word area, and this is consistent with White and colleagues (2019) that could only define the VWFA-2 in the right hemisphere in a minority of their subjects. In contrast to the left hemisphere word areas, the selectivity within these right hemisphere word areas was much more a selectivity to characters (including numbers and fake script) than to words specifically. The split-half analysis confirmed this result: we could identify word-selective voxels, but within these voxels the activity to words was not significantly higher than to the other character categories. This is consistent with previous research (Dehaene et al., 2004; Vinckier et al., 2007). One could argue about naming these areas word-selective given that they respond to all kinds of letter forms, not just letter strings of our roman alphabet but also to false fonts and to numbers. We could rarely identify the third word area, nor could we identify the pFus word area. Like in the left hemisphere, the mid-fusiform sulcus did not play a critical role in locating the word areas, but like expected it helped locate the three face areas. We replicated this face selectivity with the split-half analysis. We could define the first and often also the second hand area and this selectivity was replicated in the split-half analysis. This analysis also showed that the left hand-selective voxels were more selective to words than the right hand-selective voxels. We also found body areas separate from the hand areas and the split-half analysis showed that these body-selective voxels were significantly more active to bodies than to faces and hands (in contrast to the left hemisphere body area) and to any of the other categories of interest. In part due to the more complex nature and location of the word and hand selectivity in the right hemisphere, the hand areas could not serve as a reference point to locate the word and face selectivity. All these results from the right hemisphere are expected based on the idea that the left hemisphere is the language dominant hemisphere. Relevant to the comparison of the left and right hemisphere results, were the three left-handed participants in this study. In their left hemisphere, two of the three showed an organization just like that of the right-handed participants. The other subject showed the typical organization of the left hemisphere of the right-handed subjects in its right hemisphere instead of in its left hemisphere.

While the primary focus of this study was not to explore the functional profile of number areas, comparing them to the profile of word areas provided valuable insights into word and, more generally, symbol selectivity. Notably, the identified number areas did not exhibit selectivity specifically to numbers; instead, they demonstrated a more general selectivity to symbols, encompassing words and fake script. Thus, it could be concluded that there are areas that show selectivity to words, but also areas that are more broadly selective to symbols, and that there may not be any number areas specifically. This is an interesting observation given recent research that demonstrated the existence of number areas in the ventral occipitotemporal cortex (Cai et al., 2023). They did not however include any control category such as words (Cai et al., 2023). Combining their results with ours, the number areas that Cai and colleagues (2023) identified, may in fact not be number areas, but symbol areas. Therefore, when studying this subject of word, number and symbol selectivity, it is crucial to include at least one control category as a comparison to be able to make distinction between symbol areas in general, or more specifically word and number areas.

At last, in the representational space of the left and right hemisphere OTC, where distance between categories indicated the dissimilarity between the neural patterns, we noted that character categories formed a group separate from a group of animate and a group of inanimate categories: both numbers and fake scripts lied significantly closer to words than faces or cars (an animate and inanimate category that appeared close to words). In addition, we tested if faces lied significantly closer to words than other human-related categories as a measure of the specificity of the competition for cortical territory between words and faces. We found that faces was not the closest category to words as compared to bodies and hands, suggesting the competition may extend to human-related categories in general. This is in line with recent research from Nordt and colleagues (2021) that suggested that limbs compete with words. Furthermore, results showed that faces did not lie significantly closer to words than cars, which may signify there is no specific competition between words and faces or other human-related categories within the representational space of the occipitotemporal cortex. The competition between faces and words found in previous literature (e.g., Dehaene et al., 2010) may thus be restricted to activation to category and is not apparent in the activation patterns of different categories within the OTC. Finally, we saw in the visualization of the representational space that both hemispheres showed a very similar organization. This was in contrast with results from the functional neuroanatomy of ventral OTC, where we found a weaker word selectivity in the right hemisphere. Thus, we compared the dissimilarity between the neural patterns for words and other character categories between the left and right hemisphere. It turned out that, in contrast to what we saw in the visualization of the representational space, the patterns for words and other categories were more

similar in the right hemisphere than in the left hemisphere, which is in accordance with the results from the functional neuroanatomy of the ventral OTC.

The word forms that were used in this study (comparable to the numbers and fake script categories) were strings of letters, comparable to the fake script and numbers categories, which consisted of a string of letters of an unknown alphabet and of digits. Using these three categories, in our results of the split-half analysis, we found the word-selective areas in the left hemisphere to be significantly more selective to words than to numbers and fake script, suggesting that we at least identified areas that differentiate between letters of the typical roman alphabet versus numbers or other types of letters. Our letter strings were not pseudowords (all except one did contain at least one vowel and all were pronounceable) and we did not include any real words (aka with semantic meaning). Importantly, in the past, word selectivity has been localized with a variety of different contrasts (for a detailed overview, see Caffarra et al., 2021). For example, studies have differentiated between words, pseudowords, letter strings and individual letters (e.g., Vinckier et al., 2007), whereas others have compared text in general to false fonts or even non-linguistic categories such as faces (e.g., Cohen et al., 2002; Dehaene et al., 2010). Only studies that differentiate between several categories alike words can isolate lexical-sensitive areas (Caffarra et al., 2021). Lerma-Usabiaga and colleagues (2018) confirmed this idea: only specific types of contrast could isolate anterior from posterior word selectivity and vice versa. However, they did find that with more general contrasts (such as word-fixation) they could activate both posterior and anterior word selectivity, but of course could not make any functional distinctions between subregions in this way. Thus, in our study we identified the VWFA by a contrast between letter strings and an abundant number of other categories (some word-like and many not word-like categories). We then divided the VWFA into subregions based on clustering we found within the single-subject high-resolution anatomy of the ventral OTC. We could not make functional distinctions between the found subregions due to a lack of word-like categories and thus different contrast types. Another limitation of this study pertains to the results of the functional neuroanatomy, more specifically the organization of the areas that we identified. Across all subjects we could deduce a standard organization of word, face and hand areas in the left hemisphere, but not every subject showed this so clearly. For example, three subjects did not show the third word area in the left hemisphere. This complicated the process somewhat of the identification and localizing of areas. Supplementary Table 1 contains an overview of areas found per subject. However, small deviations from a standard organization are to be expected based on individual variability.

This study made use of 7T fMRI which was beneficial in several ways. First, it provided us with data of high quality and spatial resolution, as evidenced for example by the observation that in volume-space, activity was nicely restricted to grey matter (see Figure 5 and 6). Second, it allowed us to work

at the level of the individual, without spatial normalization or smoothing of the data. This is preferred for several reasons. Without normalization, the data stays true to the gyral and sulcal patterns of a specific brain, allowing a more accurate localization of activity (Weiner & Grill-Spector, 2013), which was a main aim of this study. Spatial smoothing may lead to averaging together of regions that actually lie distant from each other on the surface (Weiner & Grill-Spector, 2013), which would also be detrimental to the aim of our study. Group averaging can also have detrimental effects: it lacks spatial preciseness (especially when areas vary in size and/or location between subjects) and boundaries between areas can get mixed (Wandell et al., 2012; Glezer & Riesenhuber, 2013). During the creation of our analysis pipeline, we followed the advice of Brodoehl and colleagues (2020), who advised to perform a general linear model analysis and then projecting this on the individual's brain surface, without normalization or spatial smoothing. All this is in accordance with the recent call for research in the field of the VWFA to study the individual brain (Caffarra et al., 2021; Yeatman et al., 2021). Third, the use of 7T allowed us to present each subject in just one scan session with many different categories, enabling a more complete view on the functional organization of the OTC.

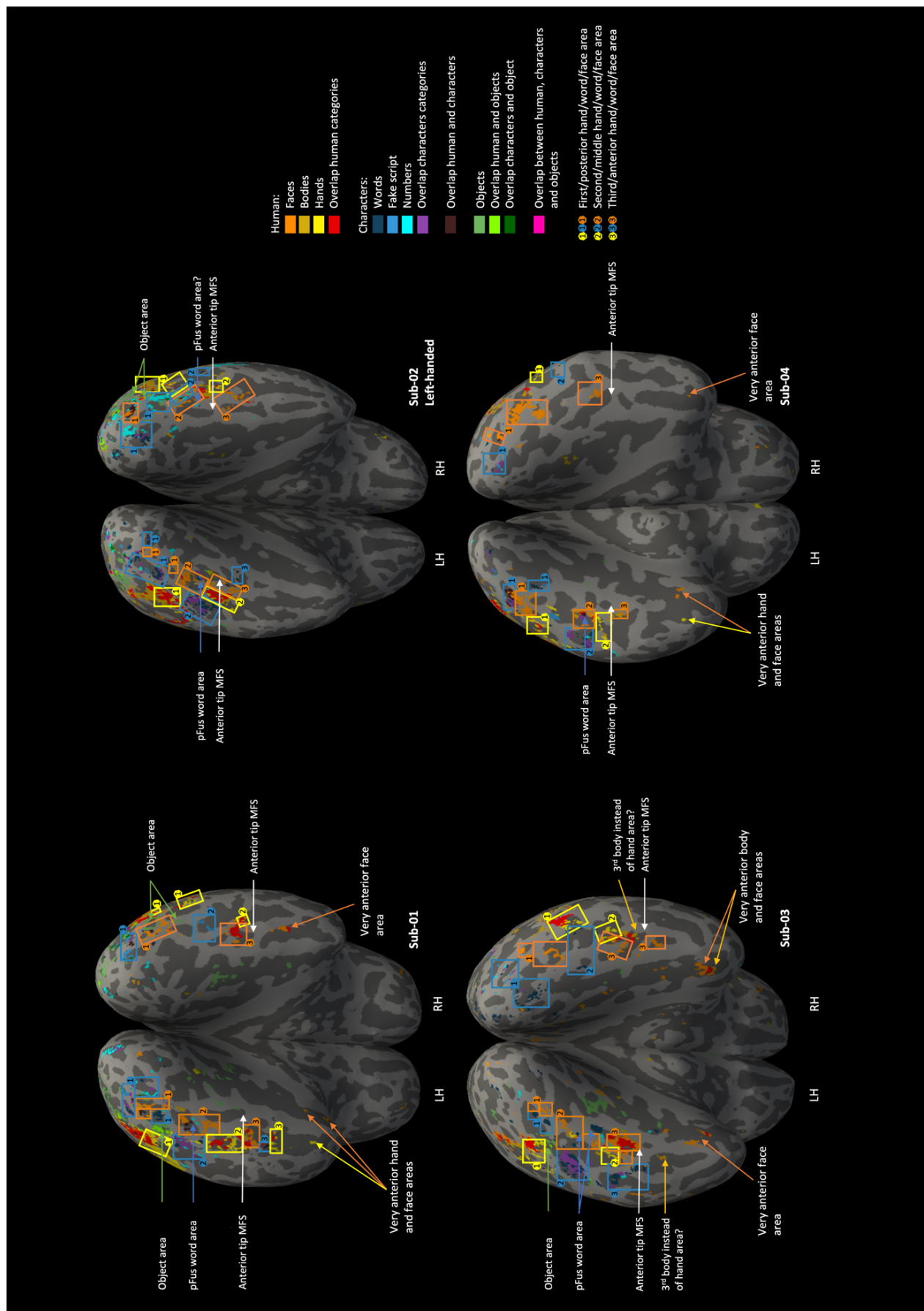
This study is the first study to our knowledge that investigates a high spatial resolution map of word selectivity and the anatomical boundaries of its subareas, together with several other possibly relevant categories. Both Caffarra and colleagues (2021) and Yeatman and colleagues (2021) recently pressed for the need to investigate where precisely, at the level of the individual brain, word selectivity was located in OTC and the possibility of subareas of word selectivity. We found three subareas of word selectivity and an extra subarea that we called the pFus word area, in consistent locations across participants in the ventral surface of OTC in the left hemisphere. These results allow future studies to define subareas of the visual word form area (VWFA) consistently. Furthermore, by relating these word areas to other areas of selectivity, we discovered hand areas that were very useful to locate these word areas and the well-known face areas (IOG, pFus and mFus), providing future studies with the option to include the hands category to be able to locate the subareas of the VWFA. We also investigated the competition between words and faces provided by the neural recycling theory. From the functional neuroanatomy that we mapped on the brain surface, we found that the hand areas provided a reference landmark for the word areas, which could suggest that other categories than just faces may be relevant to the competition with words for cortical territory. This is in line with the results found by Nordt and colleagues (2021). In addition, this study looked at the competition between words and faces within the representational space of OTC. We found no evidence that this competition for cortical territory extends to the neural activation pattern of words and faces. These results may inspire future research on the competition between words and faces as proposed originally by the neural recycling theory (Dehaene & Cohen, 2007; Dehaene et al., 2010).

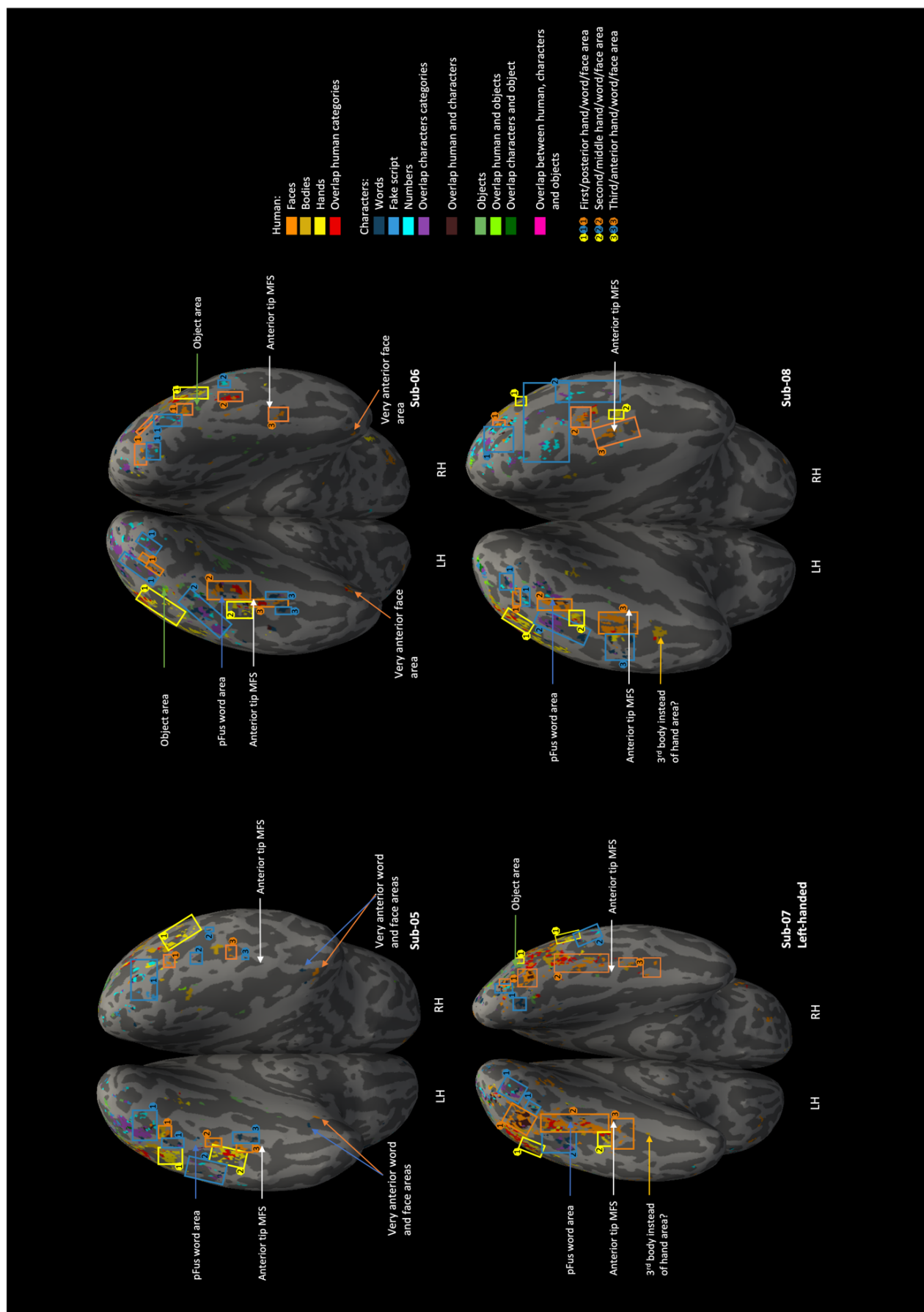
Future work could investigate differences between these subareas of the VWFA in functional connectivity with other regions of the brain. There has already been research on the relation between (subdivisions of) the VWFA and anatomical connectivity (for an overview, see Caffarra et al., 2021 and Yeatman et al., 2021), but investigating functional connectivity specifically of VWFA subareas can add new insights (Caffarra et al., 2021). In addition, future studies could investigate if and how these subareas differ in function, for example a processing gradient may very well be present going from the posterior to the anterior subarea of the VWFA. Based on many different studies (e.g., Vinckier et al., 2007; Lerma-Usabiaga et al., 2018; White et al., 2019; Woolnough et al., 2020) Caffarra and colleagues (2021) proposed a posterior to anterior processing model: VWFA can be subdivided into a posterior and anterior part, where more posterior regions represent perceptual information of written language and more anterior portions are sensitive to linguistic aspects of the presented words. The authors also considered the temporal dynamics and anatomical connectivity found in left VOTC in this word processing model. Also, a similar study that also includes pseudowords (considering the natural characteristics of language) and real words could investigate functional differences between the subregions of the VWFA as found in this study and it could add onto the results we found in representational space. Lastly, future work on the VWFA should consider that other categories than just faces, such as hands, could be interesting to include, not only when investigating the competition idea from neural recycling theories, but also to study location and/or function of (sub)areas of category selectivity.

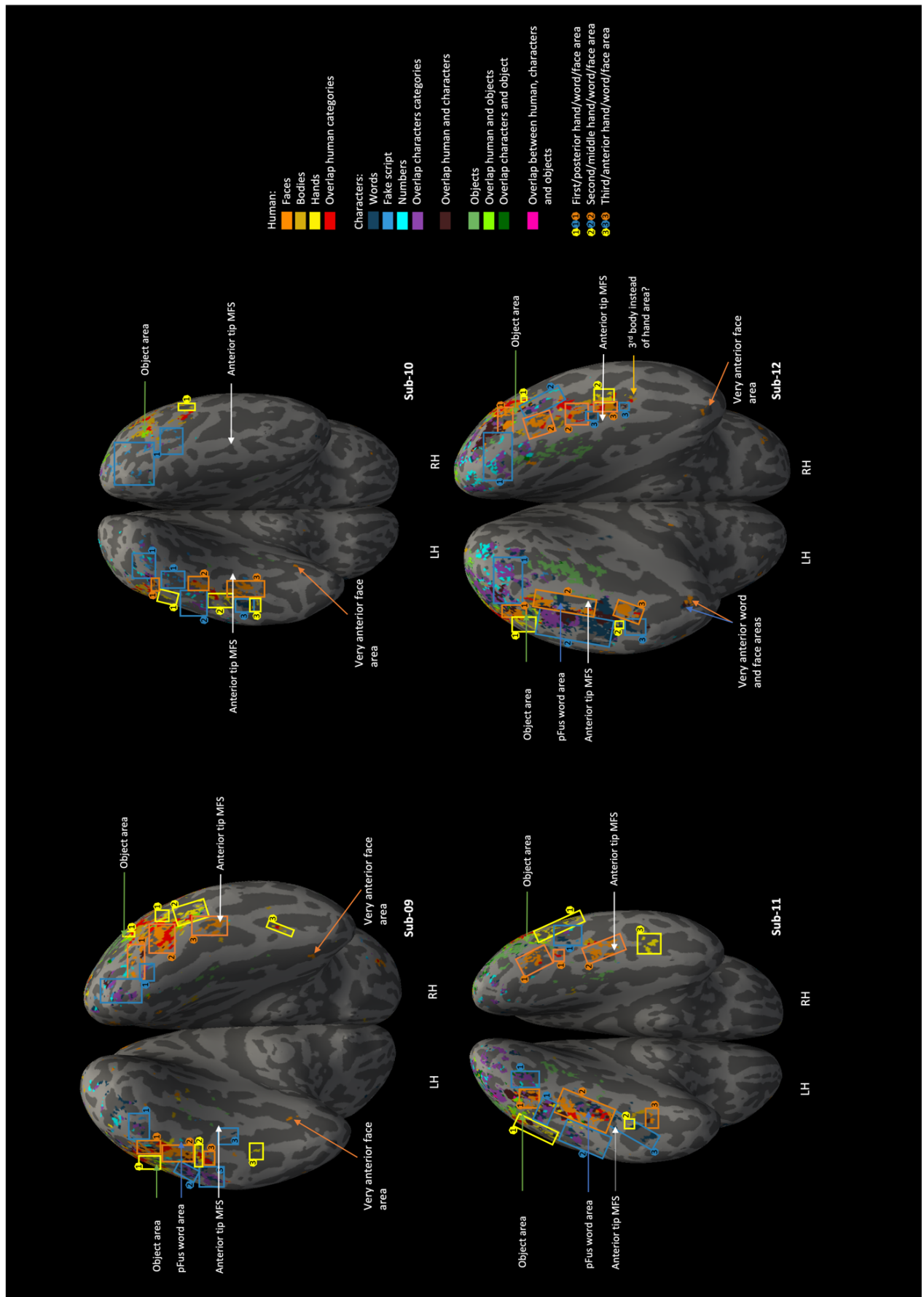
Supplementary material

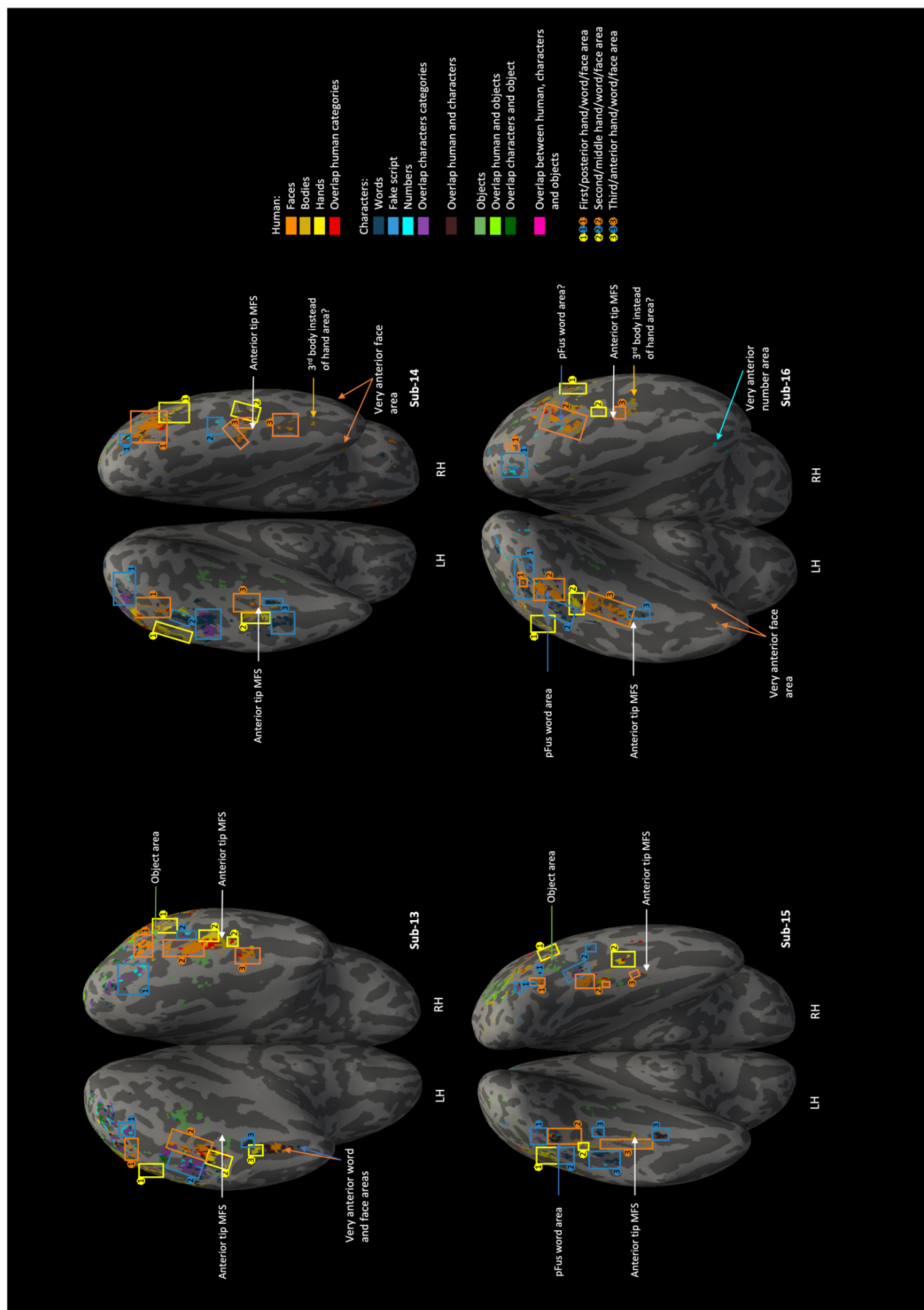
[illegible]

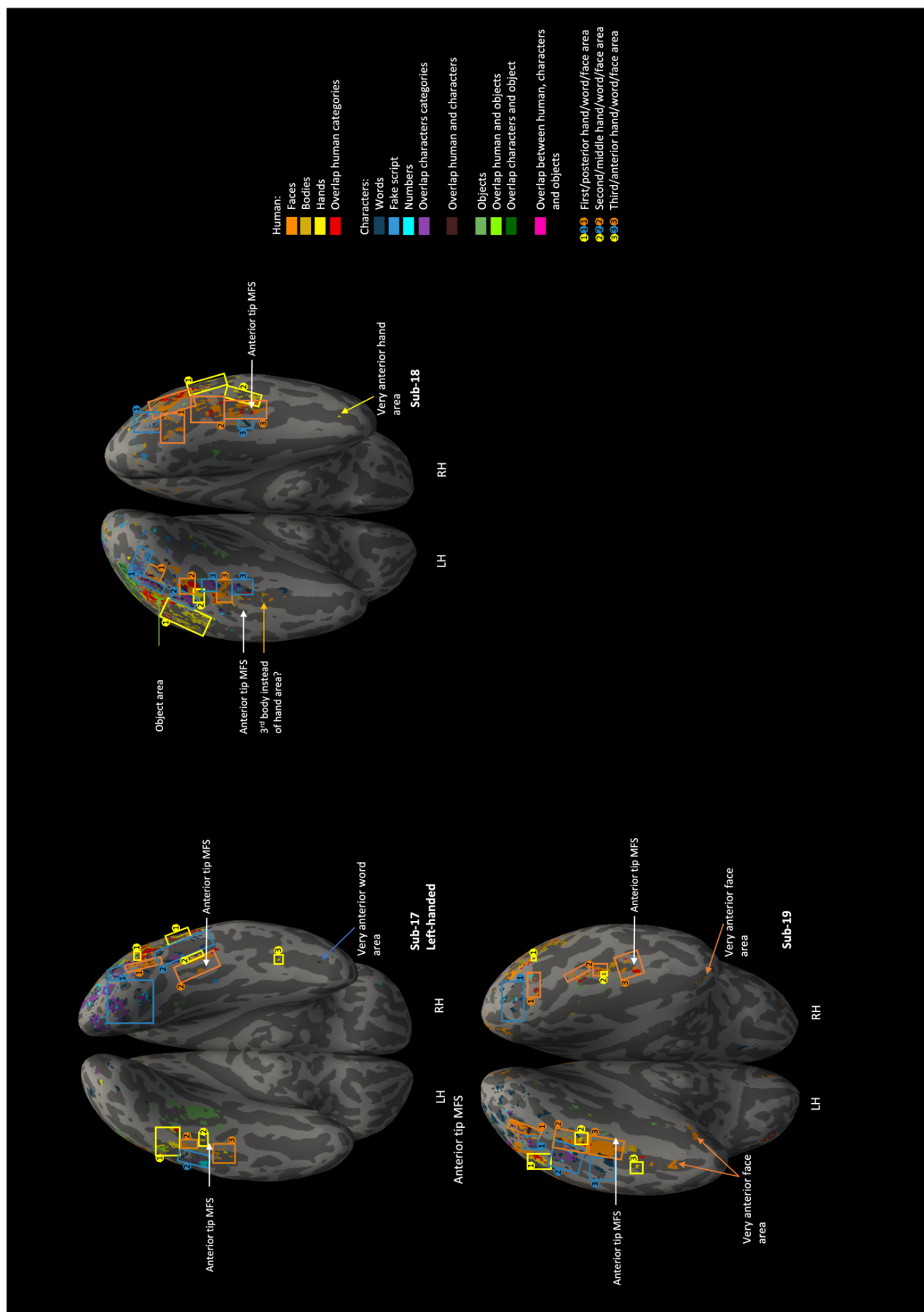
Supplementary Figure 1. Overview of which areas are found in which subjects on the ventral left (top) and in the right (bottom) hemisphere surface. Areas that are discussed on the rows of the table: the posterior word area, the posterior face area, the posterior hand area, the middle word area, the middle face area, the middle hand area, the anterior word area, the anterior face area, the anterior hand area, more anterior areas, the object region that is sometimes found posterior, the pFus word area, the ITG body area, the OTS body area. The columns are the 19 subjects and the last column indicates the total of subjects in which this area of that row could be found.



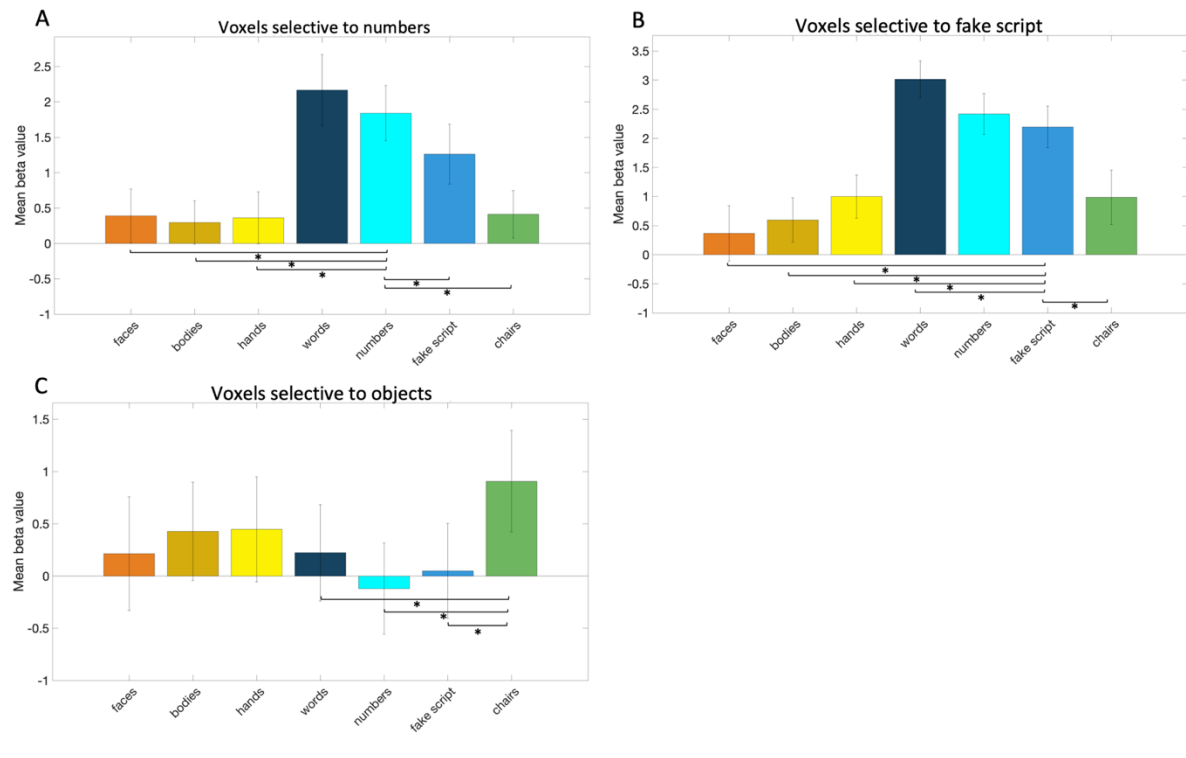




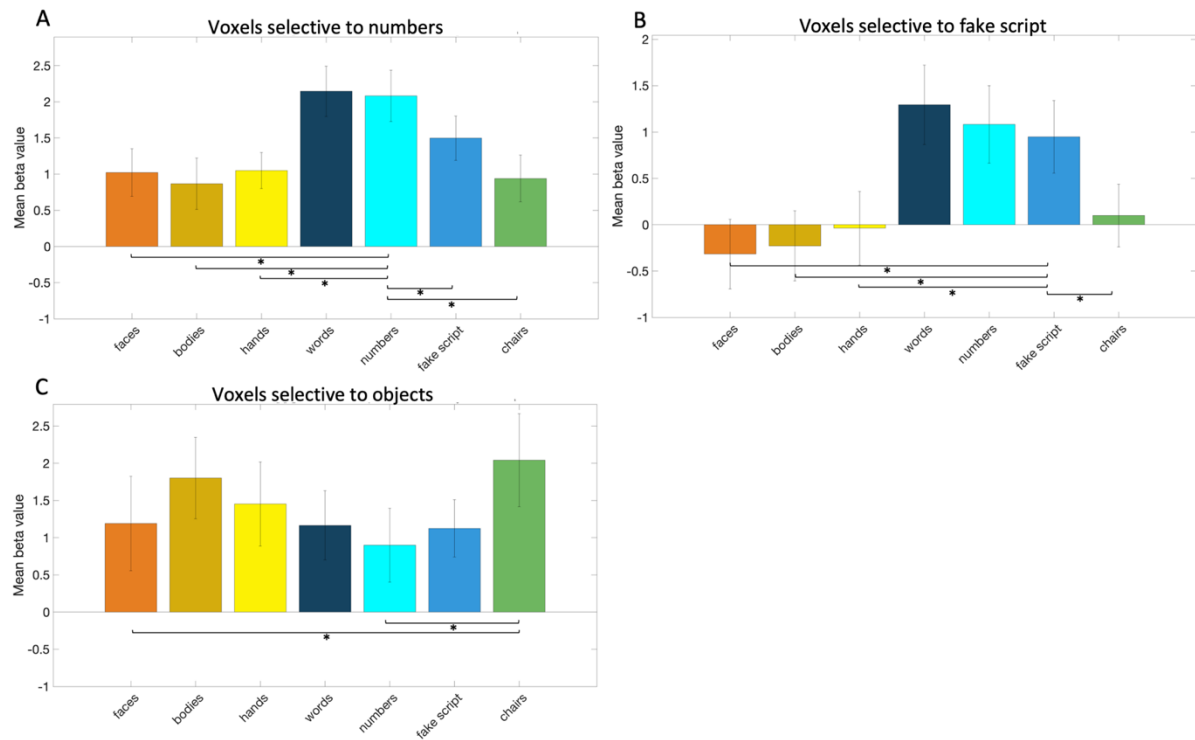




Supplementary Figure 2. Category-selective regions shown upon annotated left and right hemisphere ventral surfaces of all our participants, divided over 5 pages. Color legend on the right. Selective activation (contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected) to faces in orange, to bodies in ochre, to hands in yellow, to words in dark blue, to fake script in medium blue, to numbers in light blue, and to an objects category (chairs) in medium green. Overlap between selectivity for different human-related categories in red, overlap between selectivity for different character categories in purple, overlap between selectivity for human-related categories overlapping with character categories in brown, overlap between selectivity for human-related categories overlapping with selectivity for the objects category in light green, overlap between selectivity for character categories overlapping with the objects category in dark green, and overlap between selectivity for human-related, characters and the objects category in pink. Annotations were made using a circle, each linked to a square, with the number 1, 2 or 3 to indicate the first, second or third hand (yellow), word (dark blue), and face (orange) area. They also include several arrows that point to the posterior objects region (in green), the pFus word area (dark blue), the anterior tip of the mid-fusiform sulcus (white), a possible third body instead of the third hand area (ochre) and lastly, arrows pointing to even more anterior areas (color depends on the category). All arrows are accompanied by clarifying text. LH: left hemisphere, RH: right hemisphere.



Supplementary Figure 3. The response in number- (A), fake script- (B) and object- (chair) (C) selective voxels of the left hemisphere to seven categories (faces, bodies, hands, words, numbers, fake script and objects: chairs). Responses were calculated using data that were independent from the data used to select the voxels (see Methods). Error bars represent the standard error, and the lines and stars indicate which paired t-tests between conditions were significant ($p < .008$).



Supplementary Figure 4. The response in number- (A), fake script- (B) and object- (chair) (C) selective voxels of the right hemisphere to seven categories (faces, bodies, hands, words, numbers, fake script and objects: chairs). Responses were calculated using data that were independent from the data used to select the voxels (see Methods). Error bars represent the standard error, and the lines and stars indicate which paired t-tests between conditions were significant ($p < .008$).

Supplementary results

We investigated if the hand areas in this hemisphere also allowed to consistently locate the VWFA subareas and the typical face areas, like in the left hemisphere. We did this by first comparing the locations of these areas in the right hemisphere to the locations of the left hemisphere, and then looked at the possible guidance of the hand areas in locating the other areas. In conclusion, we found that the hand areas could not so much serve as a reference landmark like they did in the left hemisphere, in part due to the more complex nature of the word selectivity in the right hemisphere as compared to the left hemisphere. For more details, see below.

First, we investigated the posterior part of the right ventral surface of the OTC, first investigating the word selectivity. In all the subjects, we found a cluster of selectivity to some type of characters in the posterior part of the ventral surface of OTC. Compared to the left hemisphere the selectivity (whether it was to words or only to other characters) was often smaller, more distributed and in some subjects, none (or only minimal) of this selectivity was located very close to the posterior face area (IOG). Second, regarding the face selectivity (IOG), the differences between the left and the right hemisphere were minor: it appeared in the usual location and size. Third, regarding the hand selectivity, the location of the posterior hand area differed from the left hemisphere in several subjects. In subject 1, 3, 7, 10, 11 and 18, the location seemed more anterior, and this even caused it to be difficult (in subject 7 especially) to discern it from the definition of the middle hand area like we found in the left hemisphere. Lastly, in the left hemisphere in some subjects, there was a small cluster of selectivity to objects (or some kind of overlap in selectivity for objects and character or human-related categories) in between the posterior hand cluster and the posterior word and face cluster. We also found this in 10 subjects in the right hemisphere (subject 1, 2, 6, 7, 9, 10, 11, 12, 13, 15), around the posterior word and face cluster but not always exactly between those two areas and the posterior hand cluster (presumably due to the more complex location of this hand area in this right hemisphere).

Next, we investigated the middle ventral surface of the right OTC, first investigating the word selectivity. The second/middle cluster of selectivity to words was more unclear to define than the posterior one. This second word selectivity cluster appeared consistently lateral on the surface of ventral OTC in the left hemisphere and lateral from the second face cluster (pFus), but this seemed to vary in the right hemisphere. In subject 1, 2, 3, 5, 8, 14 and 15 this second word cluster was located more medial, sometimes accompanied by another small cluster that laid more lateral (subject 2, 8). Second, regarding the second cluster of selectivity to faces (pFus): this area was harder to define in the right than in the left hemisphere, because it was often less prominent. A possible reason could be the due to the presence of the second word cluster that was more often located medial in the right hemisphere, instead of lateral like in the left hemisphere. Third, regarding the hand selectivity, we

found the second hand region in many subjects, but it was harder to define in the right hemisphere compared to the left, due to the first hand area sometimes appearing in a less unexpected location (see above). Lastly, in the left hemisphere in some subjects, pFus was joined, or broken up into two, by a cluster of selectivity to words and/or other characters. This pFus word area was more difficult to discern in the right hemisphere due to pFus being absent more often or being less prominent. The often more medial location of the second word cluster also made this difficult to do. We could only find possible candidates for this definition in subject 2 and 16 but these clusters consisted only of selectivity to the numbers category and not to words.

Lastly, we investigated the anterior part of the right ventral surface of the OTC. We often found face selectivity here, but in contrast to the left hemisphere, we rarely found word or hand selectivity. In the left hemisphere we noted in some subjects an even more anterior cluster of selectivity to faces, and sometimes words. Similarly in the right hemisphere, subject 1 (although it might fit as a part of mFus too), 4, 6, 9, 12, 14 and 19 showed a cluster of selectivity to faces here. Subject 3 showed selectivity to both faces and bodies. Subject 5 showed selectivity to both faces and words. Subject 16 showed selectivity to numbers, subject 17 to words and subject 18 to hands.

We would like to conclude that for the organization of all the areas we found in the right hemisphere, the hand areas could be less of a guide to locate the other types of selectivity, because they were in another less-consistent location compared to the left hemisphere, or small (sometimes so small that it was unclear if the area was present), or the hand cluster was just not present at all (in case of the middle hand area for some subjects). In addition, the presence and location of the word clusters was also more complicated than in the left hemisphere.

References

- Abboud, S., Maidenbaum, S., Dehaene, S., & Amedi, A. (2015). A number-form area in the blind. *Nature Communications*, 6. <https://doi.org/10.1038/ncomms7026>
- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8. <https://doi.org/10.3389/fninf.2014.00014>
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron*, 21(2), 373–383. [https://doi.org/10.1016/S0896-6273\(00\)80546-2](https://doi.org/10.1016/S0896-6273(00)80546-2)
- Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1), 26–41. <https://doi.org/10.1016/j.media.2007.06.004>
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(21), 9087–9092. <https://doi.org/10.1073/pnas.0703300104>
- Barton, J. J. S., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, 58(1), 71–78. <https://doi.org/10.1212/WNL.58.1.71>
- Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2007). Differential sensitivity to words and shapes in ventral occipito-temporal cortex. *Cerebral Cortex*, 17(7), 1604–1611. <https://doi.org/10.1093/cercor/bhl071>
- Boring, M. J., Silson, E. H., Ward, M. J., Richardson, R. M., Fiez, J. A., Baker, C. I., & Ghuman, A. S. (2021). Multiple adjoining word- and face-selective regions in ventral temporal cortex exhibit distinct dynamics. *Journal of Neuroscience*, 41(29), 6314–6327. <https://doi.org/10.1523/JNEUROSCI.3234-20.2021>
- 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, 1443–1456. <https://doi.org/10.1152/jn.00619.2011>

- 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. <https://doi.org/10.1152/jn.00215.2010>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.
- Bratch, A., Vizioli, L., Yacoub, E., Engel, S. A., Burton, P. C., & Kersten, D. J. (2018). Selectivity for individual body parts within body selective visual cortex. Poster presented at OHBM 2019, Rome, Italy.
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE*, 5(5), e10773. <https://doi.org/10.1371/journal.pone.0010773>
- Brodoehl, S., Gaser, C., Dahnke, R., Witte, O. W., & Klingner, C. M. (2020). Surface-based analysis increases the specificity of cortical activation patterns and connectivity results. *Scientific Reports*, 10(1), 1–13. <https://doi.org/10.1038/s41598-020-62832-z>
- Caffarra, S., Karipidis, I. I., Yablonski, M., & Yeatman, J. D. (2021). Anatomy and physiology of word-selective visual cortex: from visual features to lexical processing. *Brain Structure and Function*, 226(9), 3051–3065. <https://doi.org/10.1007/s00429-021-02384-8>
- Cai, Y., Hofstetter, S., & Dumoulin, S. O. (2023). Nonsymbolic numerosity maps at the occipitotemporal cortex respond to symbolic numbers. *The Journal of Neuroscience*, 43(16), 2950–2959. <https://doi.org/10.1523/JNEUROSCI.0687-22.2023>
- Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: the case for the visual word form area. *NeuroImage*, 22(1), 466–476. <https://doi.org/10.1016/j.neuroimage.2003.12.049>
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, 123, 291–307. <https://doi.org/10.1093/brain/123.2.291>
- Cohen, L., Stéphane, L., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain*, 125(5), 1054–1069. <https://doi.org/10.1093/brain/awf094>
- Cohen, M. A., Alvarez, G. A., Nakayama, K., & Konkle, T. (2017). Visual search for object categories is predicted by the representational architecture of high-level visual cortex. *Journal of Neurophysiology*, 117(1), 388–402. <https://doi.org/10.1152/jn.00569.2016>
- Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., Abdi, H., & Haxby, J. V. (2012). The representation of biological classes in the human brain. *The Journal of Neuroscience*, 32(8), 2608–2618. <https://doi.org/10.1523/JNEUROSCI.5547-11.2012>

- Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, 10(4–5), 171–178. [https://doi.org/10.1002/\(SICI\)1099-1492\(199706/08\)10:4/5<171::AID-NBM453>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L)
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2), 179–94. <https://doi.org/10.1006/nimg.1998.0395>
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384–398. <https://doi.org/10.1016/j.neuron.2007.10.004>
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J.-B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *Psychological Science*, 15(5), 307–313. <https://doi.org/10.1111/j.0956-7976.2004.00674.x>
- Dehaene, S., Le Clec'h, G., Poline, J.-B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, 13(3), 321–325. <https://doi.org/10.1097/00001756-200203040-00015>
- Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Nunes Filho, G., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, 330(6009), 1359–1364. <https://doi.org/10.1126/science.1194140>
- Downing, P., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area specialized for visual processing of the human body. *Science*, 293, 2470–2473. <https://doi.org/10.1126/science.1063414>
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392, 598–601. <https://doi.org/10.1038/33402>
- Esteban, O., Blair, R., Markiewicz, C. J., Berleant, S. L., Moodie, C., Ma, F., Isik, A. I., ..., & Gorgolewski, K. J. (2018). fMRIPrep. *Software*. Zenodo. <https://doi.org/10.5281/zenodo.852659>
- Esteban, O., Markiewicz, C., Blair, R. W., Moodie, C., Isik, A. I., Aliaga, A. E., Kent, J., ..., & Gorgolewski, K. J. (2018). fMRIPrep: A robust preprocessing pipeline for functional MRI." *Nature Methods*, 16, 111–116. <https://doi.org/10.1038/s41592-018-0235-4>
- Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., & Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47, S102. [https://doi.org/10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5)
- Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., Dupont, S., Baulac, M., Dehaene, S., Adam, C., & Cohen, L. (2006). Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron*, 50(2), 191–204. <https://doi.org/10.1016/j.neuron.2006.03.031>

- Glezer, L. S., & Riesenhuber, M. (2013). Individual variability in location impacts orthographic selectivity in the “visual word form area”. *The Journal of Neuroscience*, 33(27), 11221–11226. <https://doi.org/10.1523/JNEUROSCI.5002-12.2013>
- Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of Physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance Medicine*, 44, 162–167. [https://doi.org/10.1002/1522-2594\(200007\)44:1<162::aid-mrm23>3.0.co;2-e](https://doi.org/10.1002/1522-2594(200007)44:1<162::aid-mrm23>3.0.co;2-e)
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in Python.” *Frontiers in Neuroinformatics*, 5, 13. <https://doi.org/10.3389/fninf.2011.00013>
- Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., ... Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3(1), 160044. <https://doi.org/10.1038/sdata.2016.44>
- Gorgolewski, K. J., Esteban, O., Markiewicz, C. J., Ziegler, E., Ellis, D. G., Notter, M. P., Jarecka, D., ..., & Ghosh, S. (2018). Nipype. *Software*. Zenodo. <https://doi.org/10.5281/zenodo.596855>
- Gras, V., Vignaud, A., Amadon, A., Le Bihan, D., & Boulant, N. (2017). Universal pulses: a new concept for calibration-free parallel transmission. *Magnetic Resonance in Medicine*, 77(2), 635–643. <https://doi.org/10.1002/mrm.26148>
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary- based registration. *NeuroImage*, 48(1), 63–72. <https://doi.org/10.1016/j.neuroimage.2009.06.060>.
- Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, 3(8), 837–843. <https://doi.org/10.1038/77754>
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, 15(8), 536–548. <https://doi.org/10.1038/nrn3747>
- Grill-Spector, K., Weiner, K. S., Kay, K., & Gomez, J. (2017). The functional neuroanatomy of human face perception. *Annual Review of Vision Science*, 3(1), 167–196. <https://doi.org/10.1146/annurev-vision-102016-061214>
- Harvey, A. K., Pattinson, K. T., Brooks, J. C., Mayhew, S. D., Jenkinson, M., & Wise, R. G. (2008). Brainstem functional magnetic resonance imaging: disentangling signal from physiological

- noise. *Journal of Magnetic Resonance Imaging*, 28(6), 1337–1344.
<https://doi.org/10.1002/jmri.21623>
- 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.
[https://doi.org/10.1016/S0896-6273\(02\)00662-1](https://doi.org/10.1016/S0896-6273(02)00662-1)
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of face and objects in ventral temporal cortex. *Science*, 293, 2425–2431. <https://doi.org/10.1126/science.1063736>
- Hervais-Adelman, A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P., Eisner, F., & Huettig, F. (2019). Learning to read recycles visual cortical networks without destruction. *Science Advances*, 5(9), eaax0262. <https://doi.org/10.1126/sciadv.aax0262>
- Hutton, C., Josephs, O., Stadler, J., Featherstone, E., Reid, A., Speck, O., Bernarding, J., & Weiskopf, N. (2011). The impact of physiological noise correction on fMRI at 7 T. *NeuroImage*, 57(1), 101–112. <https://doi.org/10.1016/j.neuroimage.2011.04.018>
- James, K. H., James, T. W., Jobard, G., Wong, A. C., Gauthier, I. (2005). Letter processing in the visual system: different activation patterns for single letters and strings. *Cognitive, Affective & Behavioral Neuroscience*, 5(4), 452–466. <https://doi.org/10.3758/cabn.5.4.452>
- 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), 825–41. <https://doi.org/10.1006/nimg.2002.1132>
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11), 4302–4311. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>
- Kasper, L., Bollmann, S., Diaconescu, A. O., Hutton, C., Heinzle, J., Iglesias, S., Hauser, T. U., Sebold, M., Manjaly, Z. M., Pruessmann, K. P., & Stephan, K. E. (2017). The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. *Journal of Neuroscience Methods*, 276, 56–72. <https://doi.org/10.1016/j.jneumeth.2016.10.019>
- Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., ..., & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, 13(2), e1005350. <https://doi.org/10.1371/journal.pcbi.1005350>
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1-16.
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, 74(6), 1114–1124. <https://doi.org/10.1016/j.neuron.2012.04.036>

- Kriegeskorte, N., Murm M., & Bandettini, P. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 84(4), 368–374. <https://doi.org/10.3389/neuro.06.004.2008>
- Lanczos, C. (1964). Evaluation of noisy data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 1(1), 76–85. <https://doi.org/10.1137/0701007>
- Le, R., Witthoft, N., Ben-Shachar, M., & Wandell, B. (2017). The field of view available to the ventral occipito-temporal reading circuitry. *Journal of Vision*, 17(4), 1–19. <https://doi.org/10.1167/17.4.6>
- Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences of the United States of America*, 115(42), E9981–E9990. <https://doi.org/10.1073/pnas.1803003115>
- Lerma-Usabiaga, G., Le, R., Gafni, C., Ben-Shachar, M., & Wandell, B. (2021). Interpreting sensory and cognitive signals in the cortical reading network. *BioRxiv*, 2021.09.14.460238. <https://doi.org/10.1101/2021.09.14.460238>
- Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception & Psychophysics*, 63(4), 676–697. <https://doi.org/10.3758/BF03194429>
- Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods*, 264, 47–56. <https://doi.org/10.1016/j.jneumeth.2016.03.001>
- Long, B., Störmer, V. S., & Alvarez, G. A. (2017). Mid-level perceptual features contain early cues to animacy. *Journal of Vision*, 17(6), 20. <https://doi.org/10.1167/17.6.20>
- Mattioni, S., Rezk, M., Battal, C., Bottini, R., Mendoza, K. E. C., Oosterhof, N. N., & Collignon, O. (2020). Categorical representation from sound and sight in the ventral occipito-temporal cortex of sighted and blind. *eLife*, 9, e50732. <https://doi.org/10.7554/eLife.50732>
- Moutoussis, K., & Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, 99(14), 9527–9532. <https://doi.org/10.1073/pnas.142305699>
- Nordt, M., Gomez, J., Natu, V. S., Rezai, A. A., Finzi, D., Kular, H., & Grill-Spector, K. (2021). Cortical recycling in high-level visual cortex during childhood development. *Nature Human Behaviour*, 5(12), 1686–1697. <https://doi.org/10.1038/s41562-021-01141-5>

- 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, 1–27. <https://doi.org/10.3389/fninf.2016.00027>
- Op De Beeck, H. P., 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. <https://doi.org/10.1523/JNEUROSCI.2481-06.2006>
- Orlov, T., Makin, T. R., & Zohary, E. (2010). Topographic representation of the human body in the occipitotemporal cortex. *Neuron*, 68(3), 586–600. <https://doi.org/10.1016/j.neuron.2010.09.032>
- 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. https://doi.org/10.1162/jocn_a_00411
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, 93(1), 603–608. <https://doi.org/10.1152/jn.00513.2004>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, 84, 320–341. <https://doi.org/10.1016/j.neuroimage.2013.08.048>
- Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012). Position sensitivity in the visual word form area. *Proceedings of the National Academy of Sciences of the United States of America*, 109(24), E1568–E1577. <https://doi.org/10.1073/pnas.1121304109>
- Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughhead, J., Calkins, M. E., Eickhoff, S. B., ..., & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data." *NeuroImage*, 64(1), 240–256. <https://doi.org/10.1016/j.neuroimage.2012.08.052>.
- Shum, J., Hermes, D., Foster, B. L., Dastjerdi, M., Rangarajan, V., Winawer, J., Miller, K. J., & Parvizi, J. (2013). A brain area for visual numerals. *Journal of Neuroscience*, 33(16), 6709–6715. <https://doi.org/10.1523/JNEUROSCI.4558-12.2013>
- Strother, L., Coros, A. M., Vilis, T. (2015). Visual cortical representation of whole words and hemifield-split word parts. *Journal of Cognitive Neuroscience*, 28(2), 252–260. https://doi.org/10.1162/jocn_a_00900

- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*(5761), 670–674.
<https://doi.org/10.1126/science.1119983>
- Tustison, N. J., Cook, P. A., Klein, A., Song, G., Das, S. R., Duda, J. T., Kandel, B. M., van Strien, N., Stone, J. R., Gee, J. C., & Avants, B. B. (2014). Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *NeuroImage*, *99*, 166–179.
<https://doi.org/10.1016/j.neuroimage.2014.05.044>
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word form system. *Neuron*, *55*(1), 143–156. <https://doi.org/10.1016/j.neuron.2007.05.031>
- Vogel, A. C., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2012). The putative visual word form area is functionally connected to the dorsal attention network. *Cerebral Cortex*, *22*(3), 537–549. <https://doi.org/10.1093/cercor/bhr100>
- Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, *137*, 188–200.
<https://doi.org/10.1016/j.neuroimage.2015.12.012>
- Wandell, B. A., Rauschecker, A. M., & Yeatman, J. D. (2012). Learning to see words. *Annual Review of Psychology*, *63*, 31–53. <https://doi.org/10.1146/annurev-psych-120710-100434>
- Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., Amunts, K., & Grill-Spector, K. (2014). The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage*, *84*, 453–465.
<https://doi.org/10.1016/j.neuroimage.2013.08.068>
- Weiner, K. S., & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage*, *52*(4), 1559–1573.
<https://doi.org/10.1016/j.neuroimage.2010.04.262>
- 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. <https://doi.org/10.1007/s00426-011-0392-x>
- White, A. L., Palmer, J., Boynton, G. M., & Yeatman, J. D. (2019). Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proceedings of the National Academy of Sciences*, *116*(20), 10087–10096. <https://doi.org/10.1073/pnas.1822137116>

- 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. <https://doi.org/10.3758/BRM.42.3.671>
- Wong, A. C., Jobard, G., James, K. H., James, T. W., Gauthier, I. (2009). Expertise with characters in alphabetic and nonalphabetic writing systems engage overlapping occipito-temporal areas. *Cognitive Neuropsychology*, 26(1), 111–127. <https://doi.org/10.1080/02643290802340972>
- Woolnough, O., Donos, C., Rollo, P. S., Forseth, K. J., Lakretz, Y., Crone, N. E., Fischer-Baum, S., Dehaene, S., & Tandon, N. (2020). Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway. *Nature Human Behaviour*, 5(3), 389–398. <https://doi.org/10.1038/s41562-020-00982-w>
- Yeatman, J. D., White, A. L., Yeatman, J. D., & Mall, G. (2021). Reading: the confluence of vision and language. *Annual Review of Vision Science*, 7, 487–517. <https://doi.org/10.1146/annurev-vision-093019-113509>
- Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*, 31(3), 1116–1128. <https://doi.org/10.1016/j.neuroimage.2006.01.015>
- Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1), 45–57. <https://doi.org/10.1109/42.906424>

Chapter 3: A 7T fMRI investigation of hand and tool areas in the lateral and ventral occipitotemporal cortex

This chapter as provided in this thesis is currently under review with the journal PLOS ONE.

Abstract

Previous studies demonstrated the existence of hand and tool areas in lateral and ventral occipitotemporal cortex (OTC), and an overlap between them. We reinvestigate this organization using 7T fMRI, providing us with a higher spatial resolution, a higher signal-to-noise ratio and the option to omit certain spatial preprocessing steps. All these factors could have influenced previous results, such as an artificial overlap that might not be present in the actual neural responses. In addition, we have also added several new aspects such as a comparison to faces, investigating the overlap beyond left lateral OTC, and comparisons of the similarity of neural representations of categories between several pairs of categories. First, we applied univariate analyses. There was significantly more overlap between hands and tools, compared to other overlap types in the left hemisphere of OTC. The overlap was also larger in the left lateral OTC as compared to the right lateral OTC. We found that all hand areas differentiated between tools and other types of objects, although they still preferred bodies more than tools. Regarding the tool areas, we found that they differentiated between hands and other categories like faces and objects other than tools. Left hemisphere tool areas also differentiated between hands and bodies. Second, we applied multi-voxel pattern analysis. Results indicated that in the hand areas, the neural representations between hands and tools were more similar than between hands and objects. In the tool areas, the neural representations between tools and hands and between tools and other type of objects were all equally similar.

Introduction

In both the lateral and ventral occipitotemporal cortex (OTC), areas selectively responding to specific visual categories have been identified, for example to faces (Kanwisher et al., 1997; Tsao et al., 2006), to bodies (Downing et al., 2001; Peelen & Downing, 2005), to scenes (Aguirre et al., 1998; Epstein & Kanwisher, 1998), to words (Cohen et al., 2000; Cohen et al., 2002; Baker et al., 2007), to hands (Bracci et al., 2010), and to tools (Bracci et al., 2012; Peelen et al., 2013). More specifically, the hand area, found by Bracci and colleagues (2010), was located in the lateral OTC; it could be separated from the extrastriate body area (EBA) that lied around it; and this hand area, in contrast to EBA (Downing et al., 2001), was lateralized to (i.e., responses were stronger in) the left hemisphere. In this study, they could not locate any other hand areas within the OTC that preferred hands more than bodies. However, in later research, they did find responses to hands also in right lateral OTC and in left and right ventral OTC (Bracci et al., 2016). This could be a result of the contrasts used: in their research in 2010 and 2012 they contrasted hands either with chairs (2010, 2012) or with animals (2012), whereas in the 2016 study, they contrasted hands against scrambled objects (Bracci et al., 2010, Bracci et al., 2012, Bracci et al., 2016). The authors in 2010 also noted that this hand selectivity in the left lateral OTC might be located close to previously reported responses to tools in the OTC (Chao et al., 1999; Valyear et al., 2007). This selectivity to tools was mostly lateralized to the left side of the brain (Mahon et al., 2007; Garcea & Mahon, 2014), like the responses to hands found in the lateral OTC. Within the OTC, there was also another location found to respond selectively to tools, along both the left and right medial fusiform gyrus in the ventral OTC (Mahon et al., 2007; Garcea & Mahon, 2014; Bracci et al., 2016).

It has also been suggested that a part of the hand and tool areas in the left lateral OTC overlap with each other (Bracci et al., 2012; Bracci & Peelen 2013). These responses to tools and hands could be distinguished from nearby object and motion areas and the overlap was specific to hands and tools: it did not generalize to bodies (Bracci et al., 2012). The authors in a later study also found hand and tool responses in right lateral OTC and in left and ventral OTC, that also partially overlapped (Bracci et al., 2016). This overlap within OTC, a part of the visual cortex, was a somewhat surprising observation as the categories hands and tools do not look alike visually. The organization of OTC is determined by several dimensions on which categories vary, for example in ventral OTC, selective response to animate categories appeared more lateral than to inanimate categories (Grill-Spector & Weiner, 2014). Other dimensions include for example form (Coggan et al., 2016) and manipulability (Mahon et al., 2007). None of these dimensions can explain the overlap because hands and tools differ regarding these dimensions (Knights et al., 2021). Thus, several studies attempted to understand why the hand and tool area in left lateral OTC would overlap. Several paths have been investigated, such as common

visual features (for a discussion see e.g., Peelen et al., 2013), visual imagery (for a discussion see e.g., Bracci et al., 2012) or action-related characteristics like hand-centrality and body extension (for a discussion see e.g., Bracci & Peelen, 2013), in healthy population but also in the congenital blind (e.g., Peelen et al., 2013) and people without hands (e.g., Striem-Amit et al., 2017).

In our recent 7T fMRI study (see Chapter 2 of this thesis), we could locate the hand areas found in left and right lateral and ventral OTC using a relatively stringent contrast of hands versus a variety of other object categories. This finding suggests that our dataset has the potential to re-examine the hand and tool selectivity and their overlap, avoiding several of the drawbacks of existing studies. The aforementioned studies on the overlap between hand and tool responses in lateral and ventral OTC were limited by the use of 3T fMRI and thus a standard spatial resolution (Bracci et al., 2012; Bracci et al., 2016). The overlap however might be dissociated at a higher spatial resolution (Bracci et al., 2012). As an example, face and body areas in the right fusiform gyrus were thought to overlap until they were investigated at a higher spatial resolution (Schwarzlose et al., 2005). The same argument has been made for the visual word form area by Caffarra and colleagues (2021): a region of activity might very well be separate if investigated with a high anatomical precision. In addition, studies on the overlap have significantly preprocessed the data by spatial normalization and smoothing (e.g., Bracci et al., 2012; Bracci et al., 2016). However, without such preprocessing steps, the data would be allowed to stay true to the gyral and sulcal patterns of an individual brain, enabling more accurate localizations of activity (Weiner & Grill-Spector, 2013). Spatial smoothing could lead to averaging activity together that on the surface are in truth some distance away from each other (Weiner & Grill-Spector, 2013). Therefore, we sought out to reinvestigate the hand and tool responses and their possible overlap in occipitotemporal cortex using a higher spatial resolution by using 7T fMRI and not applying such spatial preprocessing steps. In our study, we defined left and right lateral and ventral hand and tool areas and then examined their response profiles and their overlap using univariate analysis. In addition, we examined their representational space using multi-voxel pattern analysis (MVPA). We also identified face, body, and object areas as a comparison/control in our univariate analyses and included these categories (and several others) in the representational space as a reference.

Methods

This data has also been used for another study (see Chapter 2 of this thesis) and thus parts of this methods section contains previously described information.

Participants

Nineteen participants took part in this research (mean age: 30.1 ± 6.8 (23-45), sex: 11 males, 8 females). Sixteen of them were right-handed and the other three were left-handed (subject 2, 7 and 17). The participants all had normal or corrected to normal visual acuity. They all provided written informed consent. The study was approved by the ethics committee of Vrije Universiteit Amsterdam and followed the guidelines of the Declaration of Helsinki (approval number: VCWE-2020-004). Recruitment started the 1st of March 2020 and ended the 5th of September 2020.

Stimuli

Stimuli appeared on a 32" LCD screen (69.8×39.3 cm, 120 Hz) that was designed for a MRI environment (BOLDscreen, Cambridge Research Systems, UK). The resolution of the screen was 1920×1080 pixels. It was positioned at the end of the bore and participants viewed it through a mirror (distance from screen: 220 cm) that was mounted on the head coil. All stimuli were presented using MATLAB (MathWorks, Inc.) and the Psychophysics Toolbox Version 3 (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). Stimuli appeared at a semi-random location around the middle of the screen in each trial. In the middle of the screen, there was always a fixation dot present, that appeared over the stimuli whenever stimuli were present.

Twenty conditions were selected, including 19 categories and 1 scrambled control condition, so that the stimuli set would be rich and varied, according to several criteria. The conditions consisted of natural and artificial objects, familiar and new shapes (cubies and smoothies, see Op de Beeck et al., 2006), animate and inanimate shapes, objects that differ in real-world size, objects differing in how they are used and whether they are a tool or not, ... Using dimensions like these, the conditions can be grouped together in various ways. For every condition, we included different viewpoints/angles on the shapes. There was also ample variability in the identity of the stimuli in each condition. What follows is a list of all the conditions (an example image of each one is depicted in Figure 1): faces, bodies, hands, hammers, scissors, chairs, cars, musical instruments, buildings, cats, fish, flowers, trees, vegetables, cubies, smoothies, words, fake script, numbers, scrambled.

When selecting images for the conditions, we tried to minimize low-level and mid-level feature differences between the categories. For eight (bodies, buildings, cars, cats, chairs, faces, fish,

hammers) of the 20 conditions, Michael Cohen (Cohen et al., 2017) provided us with images. They ensured high within-category variability by for example, using images of objects at different angles, at different positions, and so on (Cohen et al., 2017). Like previously described, we also followed these procedures when selecting images for the other conditions. The following conditions (partially) consist of stimuli from the Bank of Standardized Stimuli (BOSS): musical instruments, flowers, vegetables, scissors (Brodeur et al., 2010). To have more stimuli for these conditions and to have stimuli for the other categories (hammers, hands, trees), we used freely available images on the internet.

We used the `randblock(rgb)` function on all the stimuli (available online for MATLAB (Mathworks, Inc.)) to create a control condition (scrambled images). We then selected 30 scrambled images and included at least one stimulus from each of the other 19 conditions. These 30 images contained more variability in the middle than at the borders. This resembles the general organization of the stimuli of the other conditions.

As new shapes, we used the cubies and smoothies categories as described in the study by Op de Beeck and colleagues (2006). The shape of both categories varied on four different dimensions. On each of the dimensions, the shape was scored from 0 to 5. We selected 16 shapes, both for cubies and for smoothies, that had an extreme position in this 4D-space (e.g., 0505, 0005) and four others that had a center position within this space (e.g. 2222, 2323). Ten of those 20 stimuli were duplicated before any further processing happened, resulting in a total of 30 stimuli. All 30 were then flipped at a semi-random angle: the duplicates were angled differently than their original version so that we would have 30 unique stimuli. We made sure that the locations of the cubies and smoothies on the grey background were not always in the center.

The words, the fake script and the numbers conditions were made of two, three or four letters/symbols. These letter/symbol strings were angled ascending or descending so that there would be difference in aspect ratio between the several stimuli of the conditions. The letters/symbols were in a bold font that was filled with a random dotted pattern. Half of the images of each condition were placed onto a white background and the others on a black background. These characteristics made these conditions more comparable to the others in terms of retinotopic envelope. The words had no semantic representation (aka they did not have a meaning), but we refer to this category as words rather than letter strings for simplicity. To create the fake script, we used a combination of two different fonts so that the letters would take on variable shapes (e.g., short, long, blocked, curvy).

On top of controlling variability within conditions in viewpoints and identity of the stimuli, some of the conditions required extra controls to match their low and mid-level features to that of the other categories. Since hammers and scissors often have an elongated shape, we translated and rotated these stimuli and made sure to include both open and closed scissors, to avoid retinotopic confounds.

For the trees, since they have a very specific shape, we used different kinds of trees and also resized them (smaller/larger). We also made sure that they were not always located on the middle section of the background but also on the upper/lower left/right. In the stimuli of the buildings, we replaced the images that included trees, with other stimuli of buildings without trees, because trees is a different condition. The stimuli of the bodies were also edited so that no hands would be visible, since this is also supposed to be a different condition.

For every condition there were 30 stimuli available. All these images were then matched for average luminance, contrast and spectral energy using the SHINE toolbox (Willenbockel et al., 2010). This procedure was inspired by the study of Cohen and colleagues (2017). We ensured that after this manipulation, the identity of the stimuli was still at least moderately clear and visible. Then, we also performed some additional controls: we created a mean greyvalue image, a standard deviation greyvalue image and an edge image per condition and then compared these images between conditions. By doing this, we ensured that the conditions did not differ on these measures.

All images are available in the following OSF project: <https://osf.io/t8jwu/>.

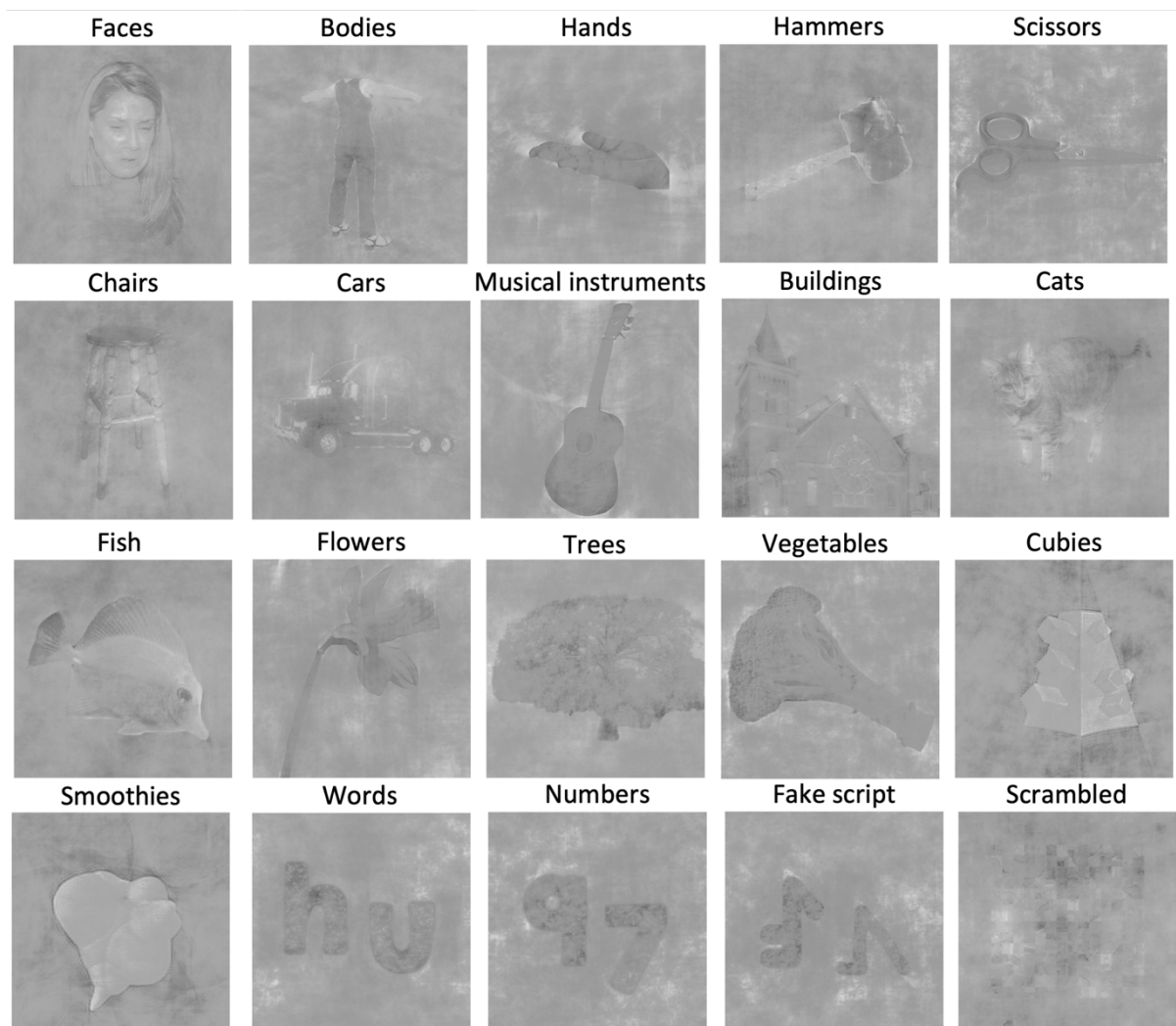


Figure 1. An example stimulus of every condition. The condition that it belongs to is written on top of the image.

Experimental design

Before entering the scanner, the participants received instructions to the task they would be doing once inside. They had to keep their focus on the fixation dot in the middle of the screen throughout the whole task. They had to perform two tasks at once. The first was a one-back task where they had to press a button with their left or right hand (this changed every subject) when an image on the screen was an identical copy of the previous image. Thus, if a stimulus was presented again but rotated or angled differently, they should not press the button because this is not an identical copy of the previous stimulus. In the second task, participants pressed a button in the other hand than the hand being used for the first task, as soon as possible when they perceived a change in category. This second task was used so category information would also be processed, and participants could not only focus on more

low-level features. To be able to perform this second task, all the categories were introduced to the participants by presenting them an example stimulus of each category.

In total, the participants went through seven runs of the experiment inside the scanner (in one subject only six, some subjects completed eight). Within a run, 40 blocks of 10.05 seconds were used. Each included 15 trials (aka stimuli). On top of those, 3 extra blocks were added (at the start, middle and end of the run) that contained only a fixation dot for 15 seconds. One trial took 0.67 seconds: during 40% of the trial (0.27 seconds) only the fixation dot was present, for the other 60% of the trial (0.40 seconds) there was a stimulus present. The stimuli in a block were always presented in a random order, and two repeats of stimuli (to create the first task) happened at random times within a block. The length of one run was 447 seconds. The first 20 experimental blocks (for the 20 conditions) were presented in a random order. After the middle fixation block, the remaining 20 blocks appeared in the order from the first 20 blocks but then reversed. We thus ended up with 140.7 seconds of data for each category in each subject with 7 runs. We used two types of runs: the 15 images for each block/condition came from the first (type 1) or the second half (type 2) of the full collection of 30 images available for that condition.

(f)MRI acquisition

Data was collected using a 7T Philips Achieva MRI scanner (Philips Healthcare, Best, The Netherlands) and an 8 channel transmit coil and a 32 channel receive coil (Nova Medical Inc, Wilmington, United States) at the Spinoza Centre for Neuroimaging in Amsterdam (the Netherlands). Universal pulses were used during scanning (for explanation see Gras et al., 2017). During scanning imaging, we used a respiration belt and a pulse oximeter on a finger of the left hand to collect respiratory and cardiac data.

To collect the functional images, we used a 3D-EPI sequence with the following parameters: volume repetition time (TR) = 1.37 s, echo time (TE) = 16.9 ms, flip angle = 13°, voxel size = 1.79x1.79x1.8 mm, field-of-view (FOV) = 200x200x176 mm and matrix size = 112 x 112 x 98. We also acquired images using an identical sequence but for reversed phase-encoding blips aka phase-encoding in the opposite direction. Those images were later used to correct for distortions. To collect anatomical images, we used a MPRAGE sequence with parameters: TR = 10 ms, TE = 3.3 ms, flip angle = 8°, spatial resolution = 0.8x0.8x0.8 mm, matrix size = 288x288x205.

Preprocessing

(f)MRI data

The data was first converted to nifti files using DCM2NIfx (Li et al., 2016) and the images were then reoriented to RAS+ with nibabel in Python. Then the data was formatted according to BIDS (Gorgolewski et al., 2016). Preprocessing was performed through fMRIPrep 20.2.0 (Esteban et al., 2018a; Esteban et al., 2018b; RRID:SCR_016216), which is based on Nipype 1.5.1 (Gorgolewski et al., 2011; Gorgolewski et al., 2018; RRID:SCR_002502). In summary, the functional images were corrected for susceptibility distortion, realigned and coregistered to the anatomical image. In addition, the brain surface was reconstructed based on the anatomical image with Freesurfer, as implemented in fMRIPrep. We did not use the spatially normalized data because we wanted to remain within the space of each participant's brain individually. This way, we could keep as much spatial and anatomical detail as possible. In its report, fMRIPrep provided details on the preprocessing steps taken and this description can be found below. To control the quality of the fMRI data, we also looked at the tSNR images (average of the timeseries divided by the standard deviation) of each participant.

There was a temporary problem with the coil during data collection: a small part of the right hemisphere is darker than the left hemisphere in some of the subjects (subject 4 to 10). In the anatomical data, such a shadow could affect the surface reconstruction. To avoid this problem, we went through several steps. For each anatomical scan, we looked at the distribution of intensity values using ITK-SNAP (Yushkevich et al., 2006) and then clipped them accordingly; then we rescaled those values using ANTs (Tustison et al., 2014); and then we performed denoising and bias correction using SPM12 and cat12. The surface reconstruction results from a fMRIPrep anat only process were then carefully checked and improved where needed. All participants' brain surfaces were successfully reconstructed. After all these steps we ran fMRIPrep in full to also perform the functional preprocessing steps while making use of these anatomical results.

The shadow did not cause trouble in the functional preprocessing. We carefully checked the preprocessing results using the fMRIPrep reports and through checking if several contrasts (e.g. faces versus all other conditions) were of similar strength and size with (1) a general linear model based on the functional space after realignment, versus (2) a general linear model after realignment, susceptibility distortion correction and coregistration to T1w space, versus (3) a general linear model after all these steps plus normalization to MNI space. This check was to ensure that all preprocessing was successful.

The following text is a detailed description of the preprocessing steps, provided by fMRIPrep. "A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al.,

2010), distributed with ANTs 2.3.3 (Avants et al. 2008; RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823; Zhang et al., 2001). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847; Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438; Klein et al., 2017). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: ICBM 152 Nonlinear Asymmetrical template version 2009c [Fonov et al., 2009; RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym]. For each of the six/seven/eight (dependent on the subject) BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. A B0-nonuniformity map (or fieldmap) was estimated based on two (or more) echo-planar imaging (EPI) references with opposing phase-encoding directions, with 3dQwarp Cox and Hyde (1997) (AFNI 20160207). Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9; Jenkinson et al., 2002). The BOLD time-series were resampled onto the following surfaces (FreeSurfer reconstruction nomenclature): fsaverage, fsnative. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD time-series were resampled into standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Several confounding time-series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following

Power (absolute sum of relative motions, Power et al., 2014) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al., 2002). FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by Power et al., 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor; Behzadi et al., 2007). Principal components are estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's aseg segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer). Many internal operations of fMRIPrep use Nilearn 0.6.2 (Abraham et al., 2014; RRID:SCR_001362), mostly within the functional processing workflow." More details on the pipeline can be found in the workflows documentation of fMRIPrep.

Physiological data

Using `scanphyslog2bids` created by Lukas Snoek (available through GitHub (<https://github.com/lukassnoek/scanphyslog2bids>), we formatted the physiological data according to BIDS (Gorgolewski et al., 2016). We corrected for physiological noise using RETROICOR (Glover et al., 2000; Hutton et al., 2011) which included using Fourier expansions of different order for the estimated phases of cardiac pulsation (3rd order), respiration (4th order) and cardio-respiratory interactions (1st order) (Harvey et al., 2008). The corresponding confound regressors were set up using the MATLAB PhysIO toolbox (Kasper et al., 2017) which is open-source code available as part of the TAPAS software collection (<https://www.translationalneuromodeling.org/tapas>). The physiological data collection did not work for all participants: participants 2, 3, 4, 5 and the first functional run of participant 18 were without physiological data.

Data analyses

Univariate analysis

With the Statistical Parametric Mapping toolbox (SPM12, Wellcome Centre for Neuroimaging, London) in MATLAB (Mathworks inc.) we created a general linear model (GLM, also called univariate analysis) for every voxel per run, for each participant. Each condition was modelled by a regressor with certain onsets and a certain duration. As a result, we had 21 regressors (20 categories + one fixation condition). Several confound regressors from fMRIPrep and the PhysIO toolbox were also included. Specifically, we used the realignment parameters and their temporal derivatives (total: 12) to minimize motion effects on the data. For each participant and run, fMRIPrep also computed a motion outliers regressors (a binary regressor for each volume in the run labeled as outlier) and these were also added to the GLM. The number of outliers varied across runs and participants. When respiratory and cardiac data were available, the resulting confound regressors (total: 18) were also added to the model. Each condition was contrasted versus all other conditions (excluding the fixation condition) to detect the voxels that respond more to that condition than to all the others. For the aims of this study, we were interested in particular in hands vs all, tools (hammers and scissors) vs all, bodies vs all, faces vs all, chairs vs all, cars vs all, musical instruments vs all and buildings vs all. To visualize these results on the brain surfaces of our participants, we selected the results of these contrasts except for the contrasts related to the different objects and buildings (the last four contrasts). For clarity of the visualization, we only selected one typical object category contrast: chairs vs all. This will be depicted by a figure where the result of these contrasts is shown on the reconstructed brain surface of example subjects. Contrasts were family-wise error (FWE) corrected and the significance level was set at $p < 0.05$.

As mentioned above, all the analyses happen within the individual brain space, thus without performing spatial normalization. This is according to the advice by Weiner and Grill-Spector (2013) as summarized in Table 2: they encourage this procedure because the data stays true to the gyral and sulcal patterns of a specific brain and this allows a more accurate localization of activity. Additionally, they also advise against spatial smoothing as this can lead to inaccurate localization of activity and to averaging of regions that are actually distant on the surface. We also follow the procedure as described by Brodoehl and colleagues (2020): for every participant, the GLM is set up based on data that is not normalized or smoothed the data in the volume-space. These results can be projected onto the brain surface of each participant separately.

Regions of interest analysis

For each participant, we visualized significant activity to faces, bodies and hands on the brain surface. We then manually drew four ROIs around this activity within the occipital and temporal cortex: a lateral left ROI, a lateral right ROI, a ventral left ROI and a ventral right ROI (see Figure 2). These four ROIs were then intersected with data from the GLM to create anatomical-functional ROIs (20 in total) using the results from five different contrasts: hands vs all, tools vs all, bodies vs all, faces vs all, and chairs vs all ($p < .05$ with FWE correction). Not all anatomical-functional ROIs could be created at this significance level for every participant.

Then, we selected the beta values of hands, hammers, scissors, bodies, faces, chairs, cars, musical instruments and buildings, per ROI and per subject, and averaged these beta values (per category) across all runs and all voxels within an ROI. As a result, we ended up with one beta value per category, per ROI and per subject. At last, we averaged across subjects to create a bar graph depicting the averaged beta value per category per ROI and calculated the standard error for each category/bar. In these bar graphs, we should not draw any conclusions about the beta value of the category that was used to create this anatomical-functional ROI.

Within each ROI, we performed several statistical tests relevant to the aims of this study. For the four hand ROIs, we tested with a paired t-test if the response to the tools category (hammers and scissors) was significantly higher than to (1) non-tool objects (chairs, cars and musical instruments but not buildings, since this is not an object), (2) bodies, and (3) faces (significance level was set at $p < .02$ based on Bonferroni correction for 3 tests, per ROI). In the tools ROIs, we tested if the response to hands was significantly higher than to (1) objects (chairs, cars and musical instruments), (2) bodies, and (3) faces (significance level was set at $p < .02$ based on Bonferroni correction for 3 tests, per ROI). In the other ROIs (faces, bodies and chairs), we compared the response to tools versus objects (chairs,

cars and musical instruments, except for chairs in the case of the chairs ROIs). The significance level was set at $p < .05$ for this test, since only this one test was performed in these ROIs.

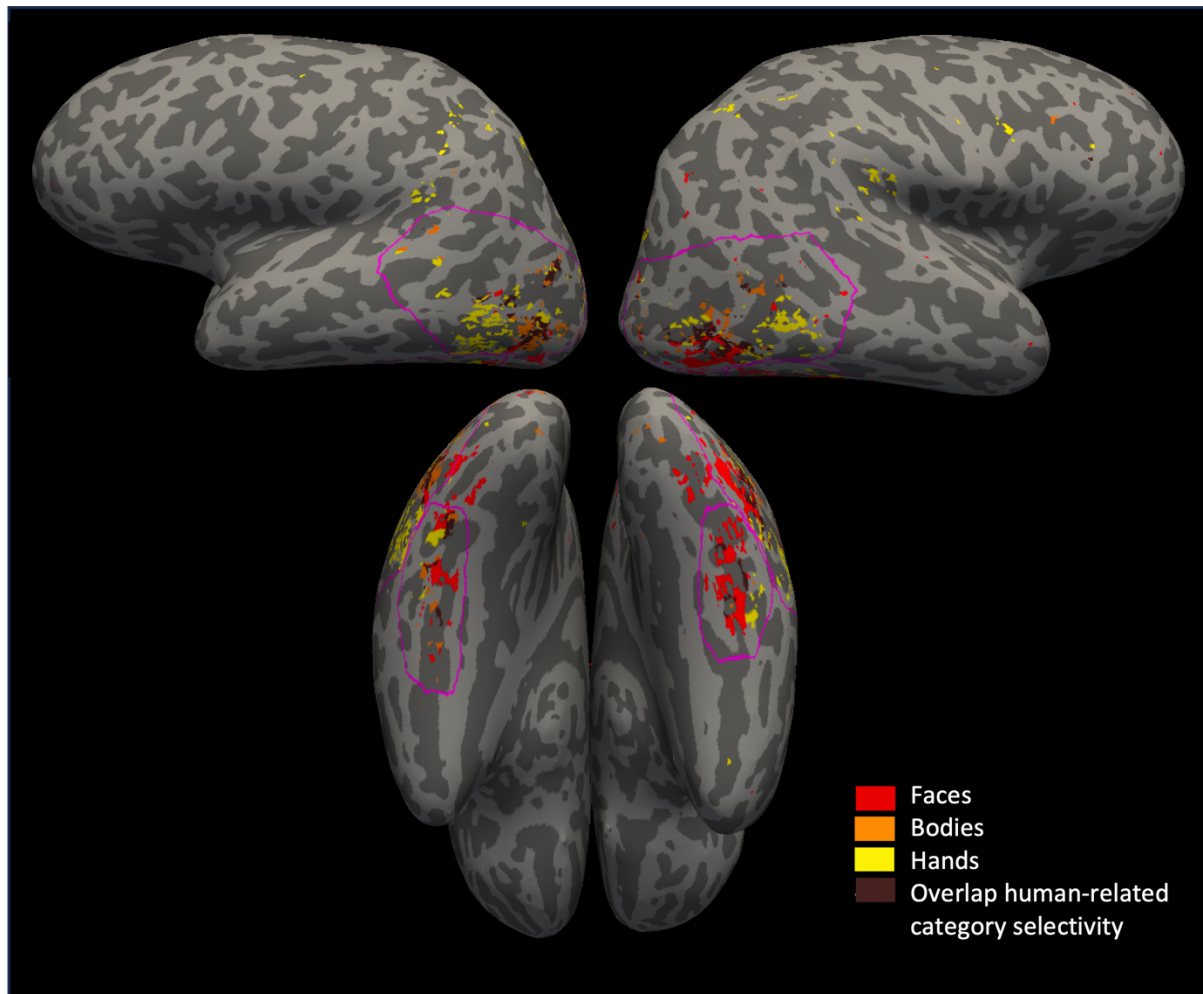


Figure 2. The anatomical regions of interest (ROI) for both the left lateral, left ventral, right lateral, and right ventral hemisphere, shown using a pink outline on a brain surface of one of the participants. Selectivity to faces shown in red on the surface, bodies in orange, hands in yellow, and any kind of selectivity overlap between these three categories in brown.

Analysis of overlap

We calculated how many voxels each anatomical-functional ROI consisted of (subjects that did not have a particular ROI were ignored in all of this analysis). In addition, we calculated how many voxels overlapped between hands and tools, between faces and tools, between bodies and tools, and between hands and objects. For each of these overlap types, we divided the number of overlapping voxels by the total of voxels in those ROIs and transformed this into a percentage. The average percentages per overlap type for lateral/ventral left/right OTC were then visualized in a bar graph.

We then performed several statistical tests. First, in left lateral, right lateral, left ventral and right ventral, we compared the percentage of overlapping voxels (again divided by the total of voxels for both categories) between hands-tools and faces-tools, between hands-tools and bodies-tools, and between hands-tools and hands-objects (significance was set at p level $< .01$ based on Bonferroni correction). Second, we also compared the percentage overlap hands-tools (divided by the total number of voxels for those two categories) between lateral and ventral in both the left and the right hemisphere and between the left and right in the lateral and ventral OTC (significance level set at $p < .01$ based on Bonferroni correction). The results of these tests are also shown on the bar graph.

Multi-voxel pattern analysis

To be able to run a multi-voxel pattern analysis within the left/right lateral/ventral hand and tool areas, we split the dataset of every participant into two. We used the anatomical ROIs drawn on the surface of every subject for the univariate analyses (see Figure 2). Then, these four ROIs per subject were intersected with the results from a GLM containing all odd runs of the dataset of that subject, for the contrast hands vs all and tools vs all (uncorrected $p < .0005$). This is similar to the univariate analyses process, except that there we intersected with a GLM based on all runs. This resulted in 8 anatomical-functional ROIs per participant. Some ROIs could not be created because there were not always voxels that reached significance.

Then, using the GLM containing the even runs for every participant, we constructed a multi-voxel pattern in response to each of the conditions for each run, by using the beta coefficient estimates of every voxel that was included in the ROI. We repeated this for each of the 8 ROIs. We used the CoSMoMVPA toolbox (Oosterhof et al., 2016) for MATLAB (Mathworks, Inc.). To decode the category for every possible pair of categories (excluding the fixation condition), for example hands-faces, tools-faces, ..., we used the cross-validated Mahalanobis distance (Walther et al., 2016), also termed the linear discriminant contrast (LDC). The code for that was written by J. Brendan Ritchie. We then saved the results in a dissimilarity matrix or in this case more specifically, a distance matrix, where each point in the matrix showed the distance between the two categories of that row and column. This distance represented a dissimilarity between the multi-voxel pattern between those two categories in that ROI. The higher the distance, the more dissimilar the patterns were. We created distance estimates for every category pair in each ROI by using a cross-validation process where the dataset is split up into training and testing partitions, according to the standard leave-one-run-out partitioner. By this process, each run was used once as a test while the other runs were used as training. As a result, there were as many distance estimates as there were runs for that participant. We then averaged across these estimates and placed the average inside the dissimilarity matrix of that participant and ROI. Every

participant matrix for each ROI was normalized by dividing all the values inside by the maximum value of that matrix. Then, per ROI, we averaged across all participants' matrices. We computed the reliability of these matrixes by correlating every participant's matrix with the average matrix (calculated excluding that subject). Finally, we performed several paired t-tests (using a Bonferroni corrected p -value) between different distances.

Multidimensional scaling (MDS) and Procrustes transformations

Multidimensional scaling (MDS) was used to visualize the main dimensions underlying the patterns in the representational dissimilarity matrices in a two-dimensional space where the distance between the points in this space was a measure for how dissimilar these points were: the higher the distance, the more dissimilar. We used the built-in MATLAB (Mathworks, Inc.) function `mdscale` with the default parameters, while minimizing the default goodness-of-fit criterion: stress and 100 replicates of the scaling. We applied MDS on the matrices, averaged across subjects, keeping left and right OTC separate. In addition, we also applied MDS on the normalized dissimilarity matrices of each individual participant. The MDS results of every subject, per ROI, were transformed to the average MDS results using a Procrustes transformation (using the built-in MATLAB (Mathworks, Inc.) function `procrustes`). We then visualized the average MDS results in a 2D space per ROI. In this space per ROI, for each category, 19 lines were drawn using the Procrustes transformed individual subject MDS position of that category. The line started in the dot of a category and ended at the coordinates (Procrustes transformed MDS results) of that category for the individual participant.

Results

The overlap between hand and tool areas

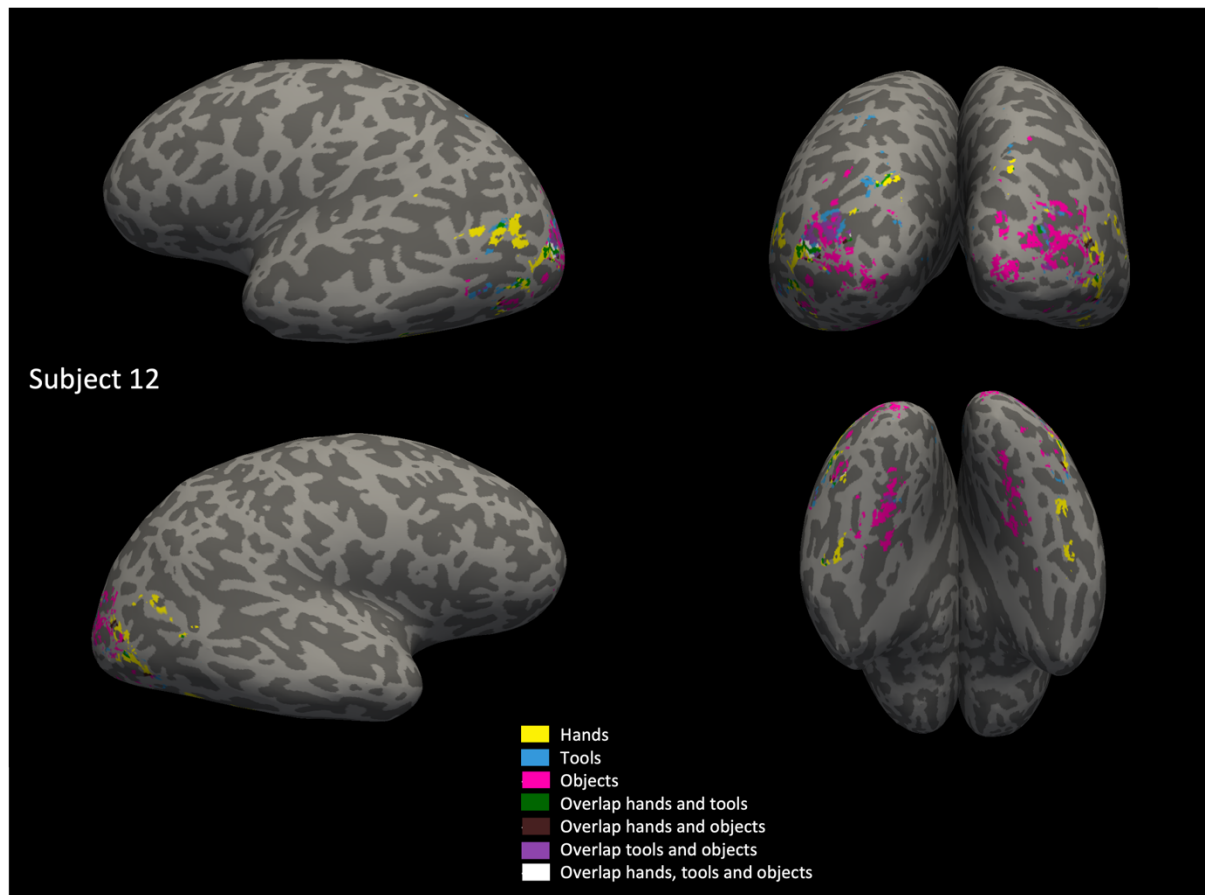


Figure 3. Hand, tool and object areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, tools in blue, objects (chairs) in pink, overlap between hands and tools in green, between hands and objects in brown, between tools and objects in purple and between hands, tools and objects in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of one example participant (color legend also on the right of the figure).

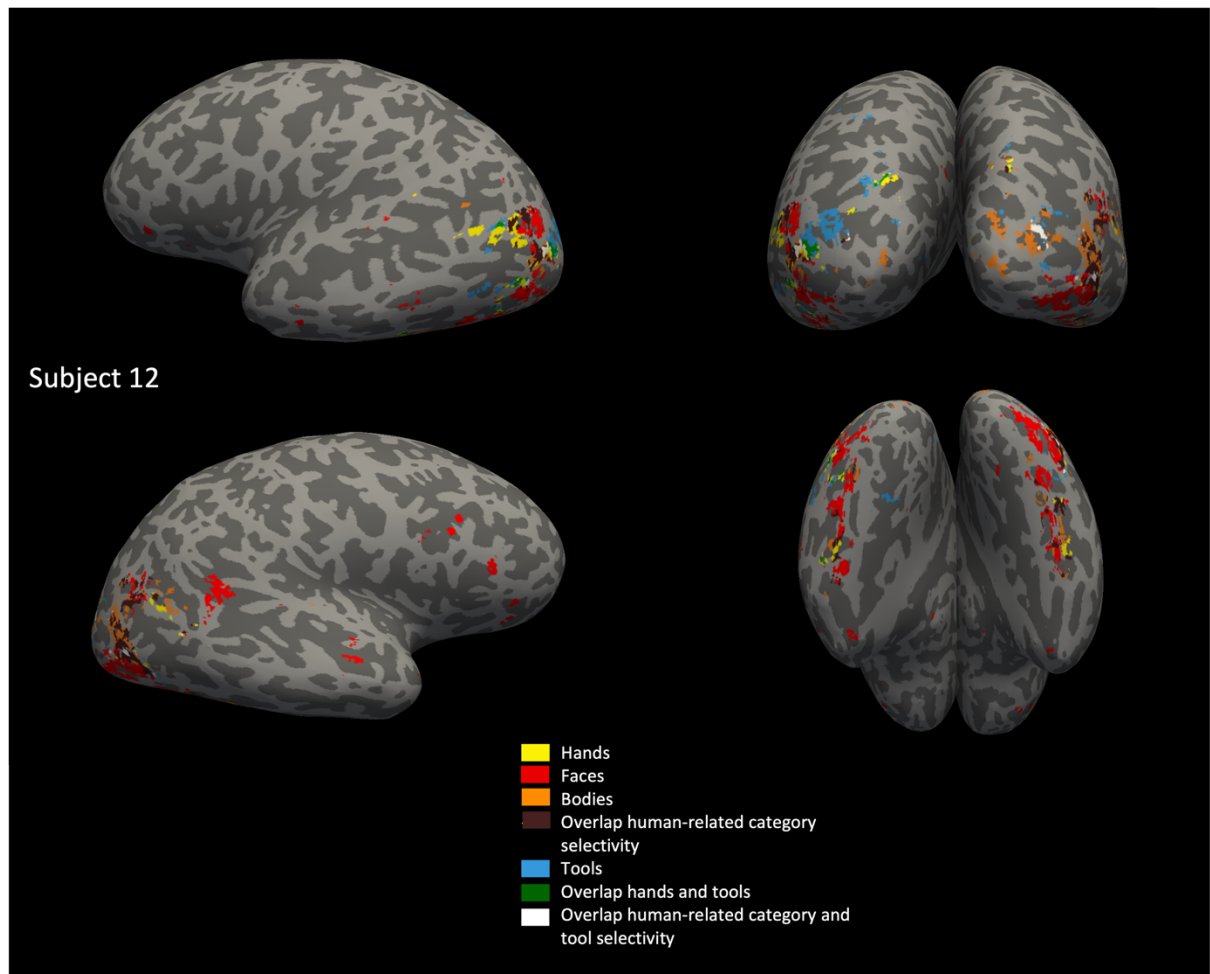


Figure 4. Hand, face, body and tool areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, faces in red, bodies in orange, overlap between some or all of these three animate categories in brown, tools in blue, overlap between hands and tools in green, and between animate categories and tools in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of one example participant (color legend also on the right of the figure).

We modelled activity to the experimental conditions using a GLM for each subject and then determined several types of category selectivity (faces vs all conditions (except the fixation condition), bodies vs all, hands vs all, tools vs all, objects: chairs vs all) on the individual's brain surface. We found selectivity to faces, bodies, hands, tools and objects in every subject (see Figure 3 and 4 for an example subject, see Supplementary Figures for more example subjects). In each anatomical ROI that we drew per subject (left/right lateral/ventral OTC), we selected only the voxels significantly ($p < .05$, FWE correction) active to hands/tools/bodies/faces/objects: chairs respectively.

We specifically analyzed the size of the overlap between selective areas, by calculating the size of different types of category overlap, normalized for total size by dividing by the total size of the two category areas in that type of overlap (results are depicted in Figure 5). We tested if the hands – tools overlap was significantly larger (i.e., $p < .01$ based on Bonferroni correction) than bodies – tools, than faces – tools, and than objects (chairs) – tools. In left lateral and in left ventral OTC, the hands – tools overlap was significantly larger than bodies – tools (left lateral: $t(17)=5.62$; $p=3.02*10^{-5}$; left ventral: $t(16)=3.16$; $p=.006$), than faces – tools (left lateral: $t(17)=8.95$; $p=7.66*10^{-8}$; left ventral: $t(16)=3.87$; $p=.001$), and than objects (chairs) – tools (left lateral: $t(17)=8$; $p=3.65*10^{-7}$; left ventral: $t(9)=3.27$; $p=.01$). Note that the degrees of freedom vary because we only include subjects that have selective voxels for both tested categories. In the right lateral OTC, the overlap hands – tools was not significantly larger than bodies – tools ($t(16)=2.47$; $p=0.02$), but it was larger than faces – tools ($t(15)=3.63$; $p=.002$) and objects – tools ($t(15)=3.17$; $p=.006$). In the right ventral OTC there were no differences, not with bodies – tools ($t(6)=0.03$; $p=.98$), with faces – tools ($t(6)=1.13$; $p=.30$) or with objects – tools ($t(2)=1$; $p=.43$). Note however that these tests are not very meaningful in right ventral OTC because we do not have sufficient selectivity to each of the individual categories. Finally, we tested if, separately in each hemisphere, the lateral overlap was larger than the ventral. Results showed no significant (i.e., $p < .01$ based on Bonferroni correction) differences ($t(16)=1.6$; $p=.13$ in the left hemisphere and $t(6)=0.8$; $p=.46$ in the right hemisphere). We also tested if there was more overlap in the left than in the right hemisphere, separately for the lateral and ventral OTC. Results showed that in the lateral OTC, there was more hands – tools overlap in the left as compared to the right hemisphere ($t(15)=4.38$, $p=.0005$), but this was not the case in the ventral OTC ($t(5)=1.8$, $p=.13$).

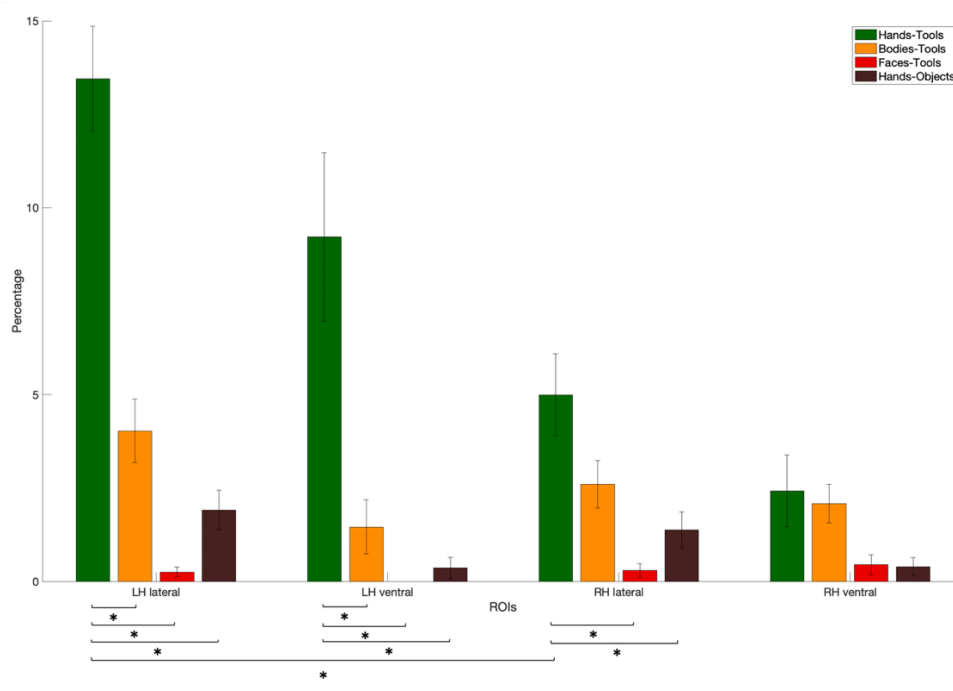


Figure 5. The average percentage of overlap between hands and tools (green), between bodies and tools (orange), between faces and tools (red) and between hands and objects (chairs, brown). Statistically significant tests between types of overlap and between different regions (left/right lateral/ventral OTC) were indicated by a *.

The response profile of hand and tool areas

In each of these anatomical-functional ROIs, we investigated the response profile to the other categories. Importantly, the results from the category should not be interpreted because that data was used to define the ROI and thus the results from that category are not independent from the definition of the ROI.

We could define the left lateral (mean size: 72.42 voxels, standard deviation: 38.67), left ventral (mean size: 9.16 voxels, standard deviation: 6.13) and right lateral hand area (mean size: 41.26 voxels, standard deviation: 34.33) in all 19 subjects. The right ventral area (mean size: 5.26 voxels, standard deviation: 8.02) could be defined in 15 subjects. In the left lateral hand area (see Figure 6), we found that the activity in response to the two tools categories (hammers and scissors) was significantly (i.e., $p < .02$ based on Bonferroni correction) higher than to faces ($t(18)=7.4$; $p=7.35 \times 10^{-7}$) and to other objects (chairs, cars and musical instruments, $t(18)=3.93$; $p=.001$). However, the response to bodies was higher than to the tools ($t(18)=-3.91$; $p=.001$). Similar results were found in the right lateral hand area: tools – faces ($t(18)=2.71$; $p=.01$), tools – objects ($t(18)=6.54$; $p=3.82 \times 10^{-6}$), tools – bodies ($t(18)=-6.03$; $p=1.06 \times 10^{-5}$). In the left ventral hand area, the response to tools did not differ significantly from faces ($t(18)=1.09$; $p=.29$) or from bodies ($t(18)=-1.03$; $p=.2$), but it was higher than to objects

($t(18)=3.54$; $p=.002$). In the right ventral hand area, the response to tools did not differ from faces ($t(14)=1.16$; $p=.27$), but it was lower than bodies ($t(14)=-4.33$; $p=.0007$) and higher than objects ($t(14)=4.02$; $p=.001$). Thus, all hand areas differentiated between tools and other objects (that included chairs, cars and musical instruments), but the response to bodies (except in the left ventral hand area) was higher than to tools. Interestingly, in the lateral hand areas, we did find evidence that the tools were preferred over faces, which is somewhat surprising because faces is an animate category (like hands).

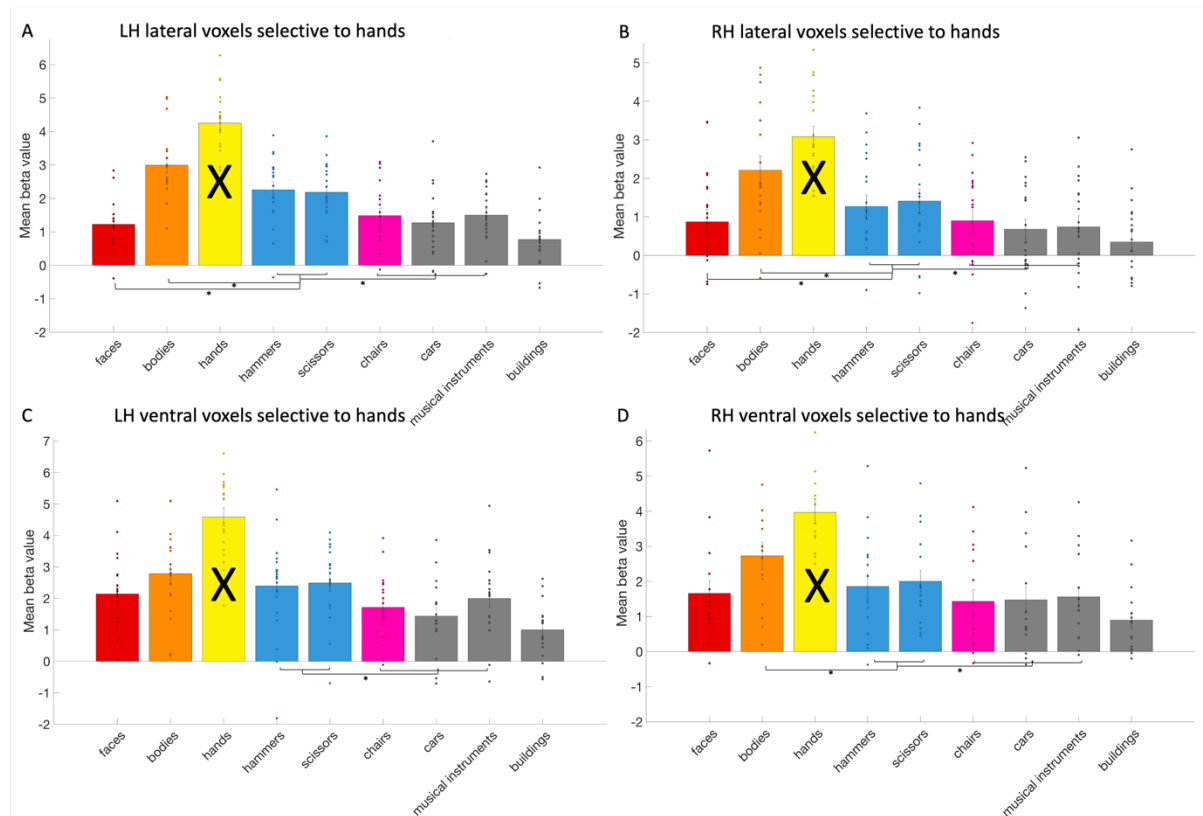


Figure 6. The response profile on average and per subject (visualized by a dot per bar thus category) to faces, bodies, hands, hammers, scissors, chairs, cars, musical instruments, and buildings in the (A) left lateral (B) right lateral (C) left ventral and (D) right ventral hand areas. The results regarding to hands could not be interpreted (indicated by a black cross across the bar) as this was already used to define the hand areas. Significant statistical tests between categories were indicated using a *.

We could define the left lateral tool area (mean size: 34.32 voxels, standard deviation: 22.65) in 18 subjects, the left ventral tool area (mean size: 6.68 voxels, standard deviation: 8.73) in 17 subjects, the right lateral tool area (mean size: 13.53 voxels, standard deviation: 15.89) in 17 subjects, and the right ventral tool area (mean size: 2.68 voxels, standard deviation: 5.79) in 8 subjects. In the left lateral tool area (see Figure 7), we found that the response to hands was significantly (i.e., $p < .02$ based on

Bonferroni correction) higher than to faces ($t(17)=12.78$; $p=3.82 \cdot 10^{-10}$), bodies ($t(17)=6.94$; $p=2.38 \cdot 10^{-6}$) and objects (chairs, cars and musical instruments, $t(17)=8.78$; $p=1.01 \cdot 10^{-7}$). In the right lateral tool area, results were similar except that the response to hands and bodies ($t(16)=0.08$; $p=.94$) did not differ significantly, but the response to hands was still higher than to faces ($t(16)=5.06$; $p=.0001$) and objects ($t(16)=3.4$; $p=.004$). In the left ventral tool area, results were the same as in the left lateral tool area: the response to hands was significantly higher than to faces ($t(16)=7.07$; $p=2.58 \cdot 10^{-6}$), bodies ($t(16)=4.7$; $p=.0002$) and objects ($t(16)=3.65$; $p=.002$). In the right ventral tool area, hands scored higher than faces ($t(7)=4.95$; $p=.002$), but not higher than bodies ($t(7)=0.85$; $p=.42$) or objects ($t(7)=1.77$; $p=.12$). Thus, the tool areas in the left hemisphere preferred hands over the other categories (faces, bodies, objects that included chairs, cars and musical instruments). In the right hemisphere, this was only the case to some extent, hands were differentiated from faces and in the case of the lateral area, also from objects, but not from bodies.

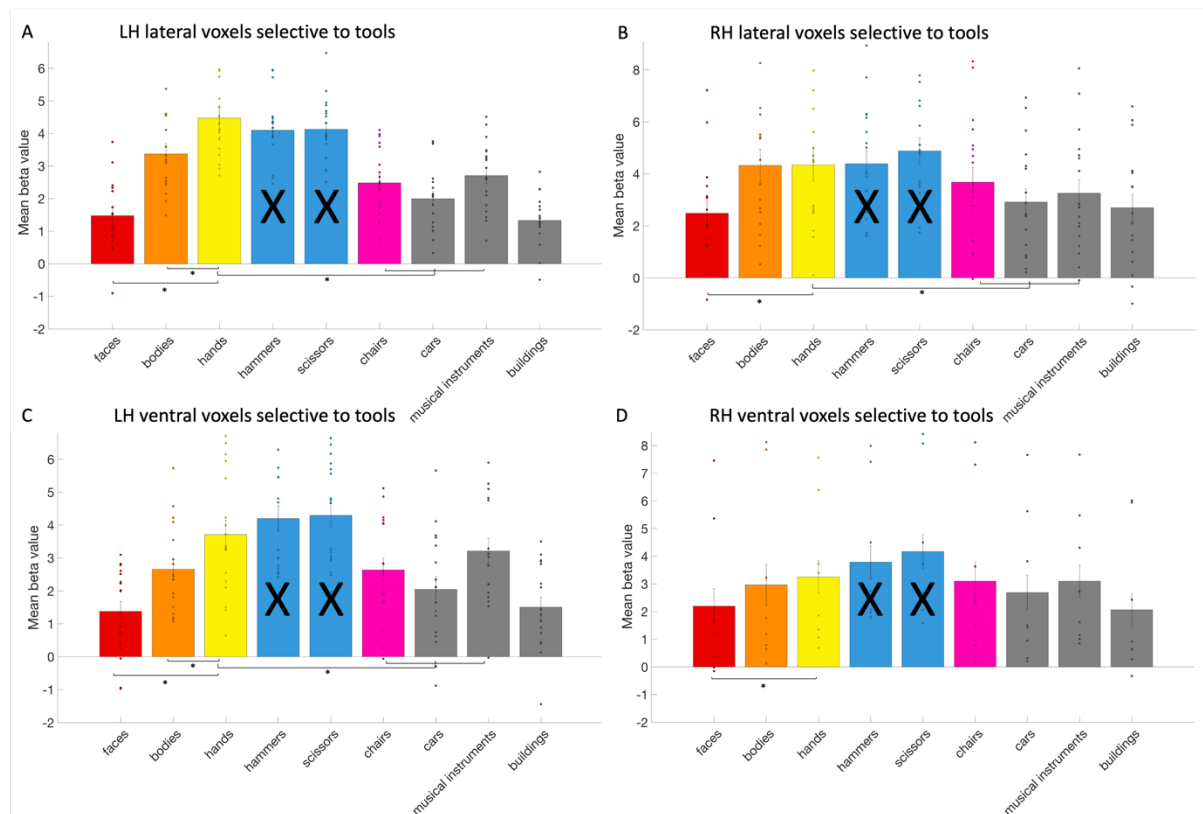


Figure 7. The response profile on average and per subject (visualized by a dot per bar thus category) to faces, bodies, hands, hammers, scissors, chairs, cars, musical instruments, and buildings in the (A) left lateral (B) right lateral (C) left ventral and (D) right ventral tool areas. The results regarding to tools could not be interpreted (indicated by a black cross across the bar) as this was already used to define the tool areas. Significant statistical tests between categories were indicated using a *.

We could define the left lateral (mean size: 56.42 voxels, standard deviation: 31.5), left ventral (mean size: 8.05 voxels, standard deviation: 5.7) and right lateral body area (mean size: 73 voxels, standard deviation: 45.68) in all 19 subjects and the right ventral body area (mean size: 10.63 voxels, standard deviation: 10.9) in 17 subjects. In the body areas (see Figure 8), as a control compared to the hand areas, we tested only if these body areas could differentiate between tools and other objects. This was only the case in the right lateral body area, where the response to tools was significantly (i.e., $p < .05$) higher than to other objects ($t(18)=3.15$; $p=.006$). The results for the left lateral body area for tools – objects was: $t(18)=0.87$; $p=.4$; for the left ventral body area: $t(18)=-0.28$; $p=.78$; for the right ventral body area: $t(16)=0.21$; $p=.84$. Thus, only the right lateral body area preferred tools more than other objects (including chairs, cars and musical instruments).

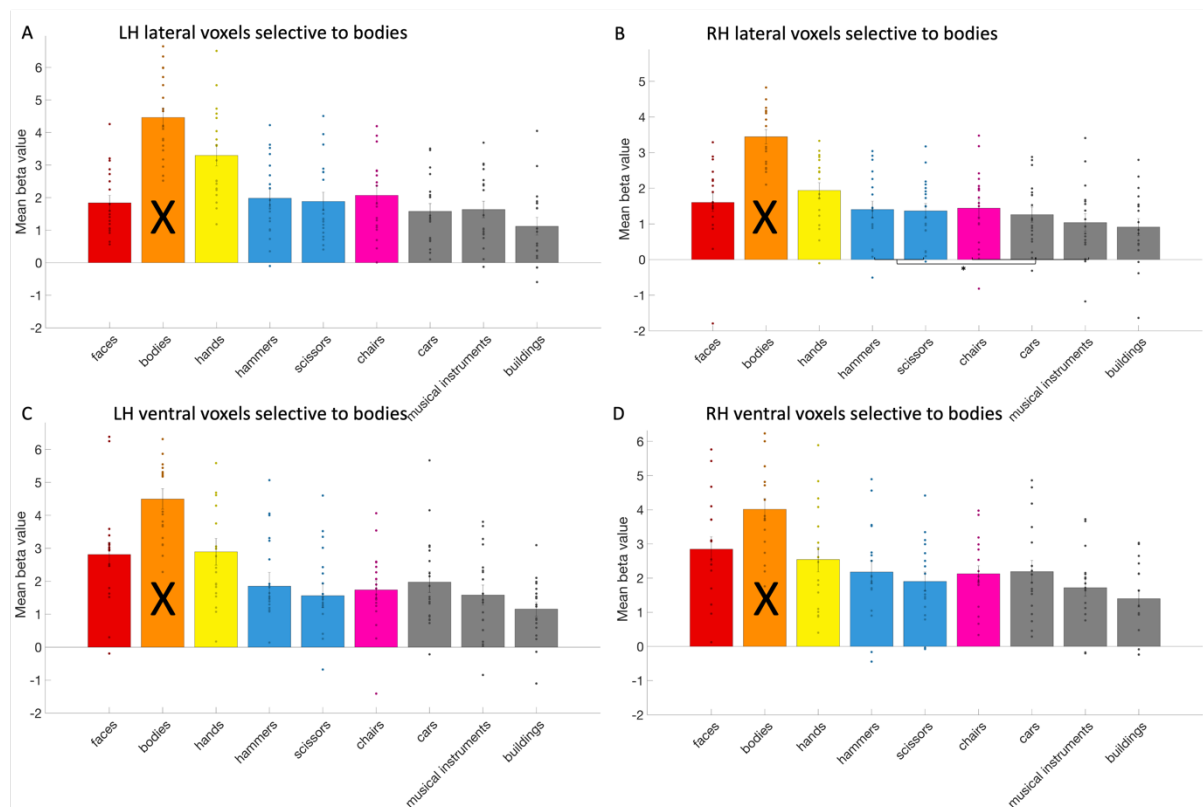


Figure 8. The response profile on average and per subject (visualized by a dot per bar thus category) to faces, bodies, hands, hammers, scissors, chairs, cars, musical instruments, and buildings in the (A) left lateral (B) right lateral (C) left ventral and (D) right ventral body areas. The results regarding to bodies could not be interpreted (indicated by a black cross across the bar) as this was already used to define the body areas. Significant statistical tests between categories were indicated using a *.

We could define the left lateral (mean size: 37 voxels, standard deviation: 36.87) and ventral face area (mean size: 29.21 voxels, standard deviation: 21.15) in 19 subjects, the right lateral (mean size:

40.89 voxels, standard deviation: 39.63) in 18 subjects and the right ventral (mean size: 17.26 voxels, standard deviation: 12.85) in 17 subjects. Just like for the body areas, we tested if the face areas (see Figure 9), as a control for the hand areas, could differentiate between tools and other objects. This was only the case in the right ventral face area, where the response to objects was significantly (i.e., $p < .05$) higher than to tools ($t(16)=-2.57$; $p=.02$). The results for the left lateral face area for tools – objects was: $t(18)=-1.32$; $p=.2$; for the right lateral face area: $t(17)=-0.57$; $p=.58$; for the left ventral face area: $t(18)=-2.01$; $p=.06$. Thus, only the right ventral face area preferred objects (including chairs, cars and musical instruments) more than tools, which is the opposite direction of the tool preference observed in the hand areas.

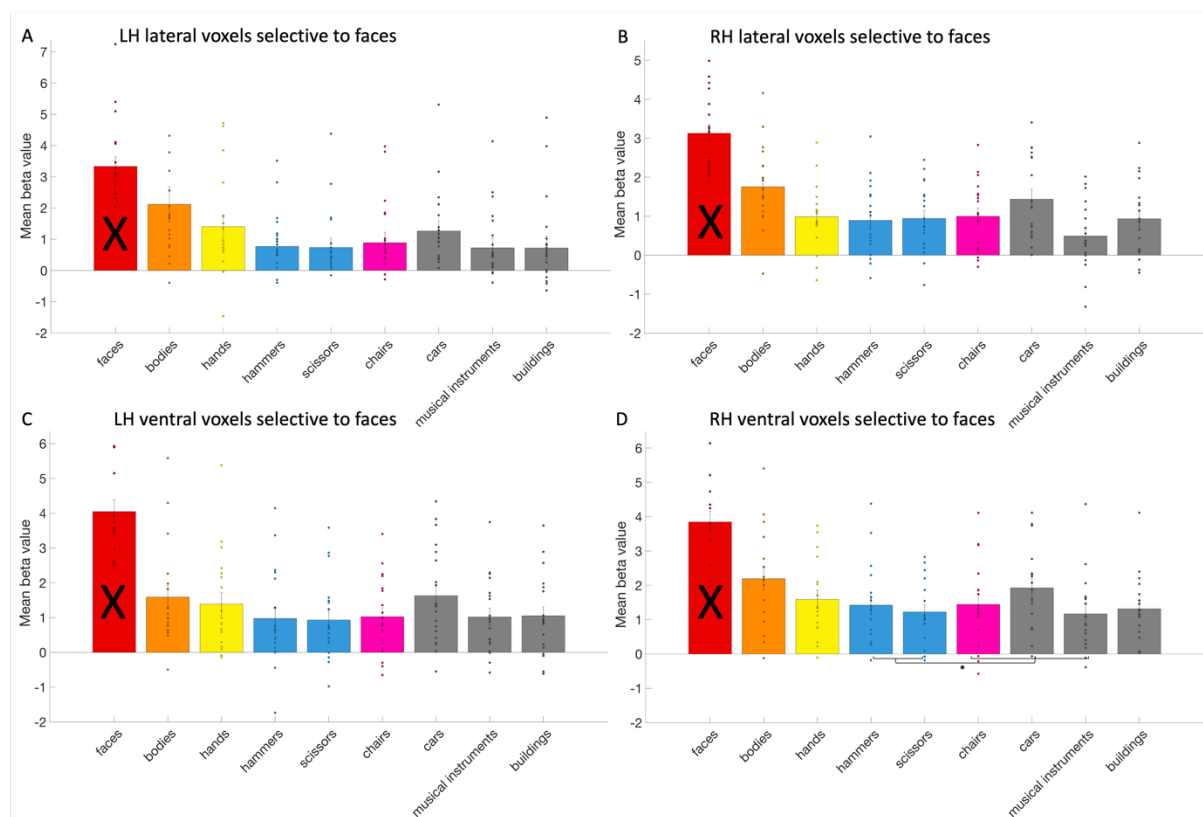


Figure 9. The response profile on average and per subject (visualized by a dot per bar thus category) to faces, bodies, hands, hammers, scissors, chairs, cars, musical instruments, and buildings in the (A) left lateral (B) right lateral (C) left ventral and (D) right ventral face areas. The results regarding to faces could not be interpreted (indicated by a black cross across the bar) as this was already used to define the face areas. Significant statistical tests between categories were indicated using a *.

As a last control, we identified object areas (see Figure 10) using the chairs category. We could define a left lateral object area (mean size: 34.21 voxels, standard deviation: 43.06) in 19 subjects, a left ventral (mean size: 1.47 voxels, standard deviation: 2.2) in 12 subjects, a right lateral (mean size:

28.32 voxels, standard deviation: 31.51) in 18 subjects and a right ventral (mean size: 1.47 voxels, standard deviation: 3.06) in 8 subjects. We then tested in each of them if they differentiated between tools and other objects (cars and musical instruments). We found such a differentiation in the lateral areas, where objects scored higher than tools (left lateral: $t(18)=-2.92$; $p=.009$; right lateral: $t(17)=-2.72$; $p=.01$). There was only a trend in left ventral object area ($t(11)=-2.07$; $p=.06$) and in the right ventral object area ($t(7)=-2.1$; $p=.07$). Thus, generally (whether significant or a trend), object/chair areas preferred other objects more than tools, again in contrast to the preference for tools found in the hand areas.

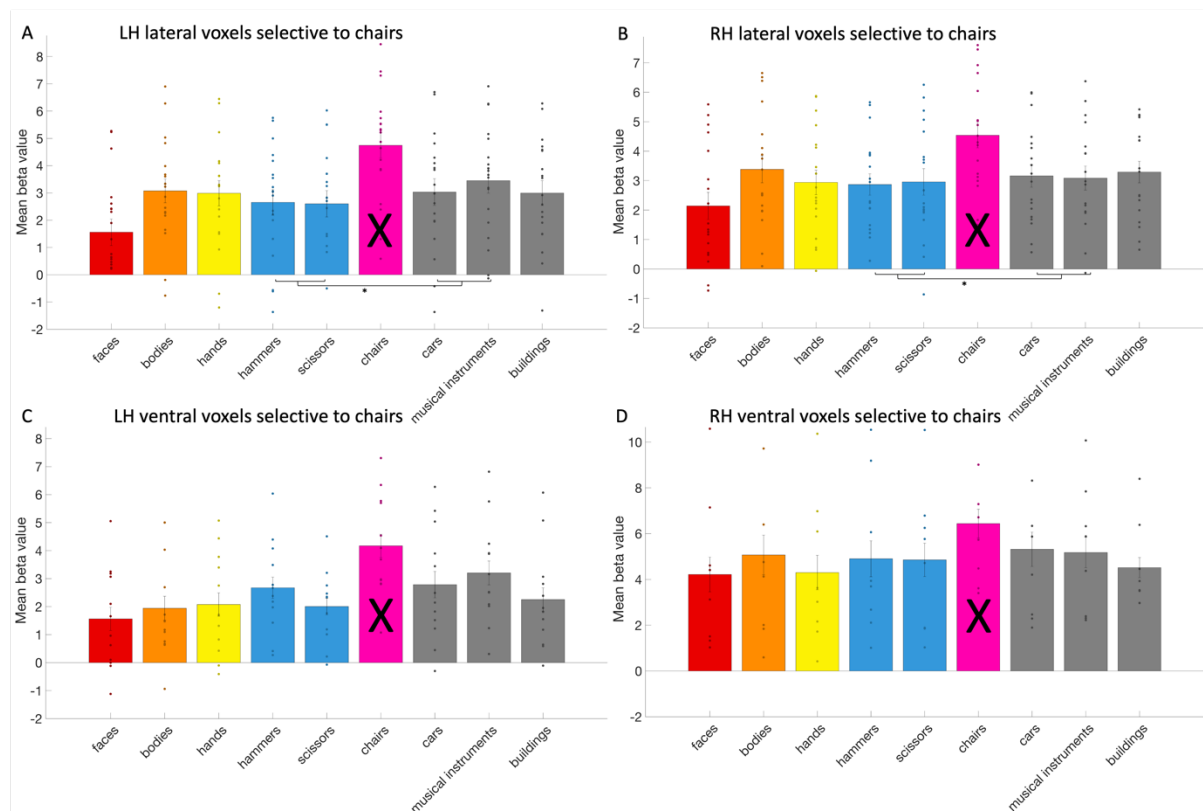


Figure 10. The response profile on average and per subject (visualized by a dot per bar thus category) to faces, bodies, hands, hammers, scissors, chairs, cars, musical instruments, and buildings in the (A) left lateral (B) right lateral (C) left ventral and (D) right ventral object/chair areas. The results regarding to the object chairs could not be interpreted (indicated by a black cross across the bar) as this was already used to define the object/chair areas. Significant statistical tests between categories were indicated using a *.

The representational space of hand and tool areas

We also mapped the representational space in response to all categories of the hand and tool areas in left/right lateral/ventral OTC. To achieve this, we defined the ROIs by intersecting the

anatomical ROIs that we drew with the results of a GLM (hands vs all or tools vs all, except fixation) based on only half of the runs, for every subject. We then ran a multi-voxel pattern analysis using independent data (the GLM based on the other half of the runs), per subject. This resulted in a matrix per ROI and per subject, that were consequently normalized and then averaged across participants. To estimate reliability, each subject's matrix was correlated to the average matrix. The mean correlation was the following for the ROIs: left lateral hand area: $r = .68$, left ventral hand area: $r = .45$, right lateral hand area: $r = .51$, right ventral hand area: $r = .42$, left lateral tool area: $r = .58$, left ventral tool area: $r = .35$, right lateral tool area: $r = .46$, right ventral tool area: $r = .29$. We then visualized the MVPA results of every ROI in a two-dimensional space using MDS (see Figure 11 for the hand areas and Figure 12 for the tool areas). In this 2D space, categories that lay closely together showed more similar multi-voxel patterns (neural representations) than categories that lay more apart. Inside these 2D spaces, we also showed the MDS results of every subject after Procrustes transformation: one line for every subject, per category and per space.

We first focused on results from the hand areas (see Figure 11). We could define (the size of the ROI needed to be larger than 1 voxel) a left lateral hand area in all 19 subjects (mean size: 125.16 voxels, standard deviation: 73.97), a left ventral in all 19 subjects (mean size: 23.37 voxels, standard deviation: 14.24), a right lateral in all 19 subjects (mean size: 78.58 voxels, standard deviation: 86.07) and a right ventral (mean size: 9.21 voxels, standard deviation: 8.02) in 15 subjects.

Our main interest, as in the univariate analysis was the possible differentiation between tools and other objects (as before, we chose chairs, cars and musical instruments as typical objects to compare with) by the hand areas. In each hand area, we thus tested (using a paired t-test), if the distance as noted in the dissimilarity matrix (the normalized LDC value of each subject) of hands – hammers and hands – scissors was smaller than the distance between hands – chairs, hands – cars, and hands – musical instruments. Indeed, in all hand areas except the right ventral hand area (but it did show a trend), the distance was significantly (i.e., $p < .01$ based on Bonferroni correction) smaller (left lateral: $t(18) = -11.26$, $p = 1.39 \times 10^{-9}$; left ventral: $t(18) = -3.8$, $p = .001$; right lateral: $t(18) = -9.82$, $p = 1.19 \times 10^{-8}$; right ventral: $t(14) = -2.12$, $p = .05$). This signified that the neural representations of hammers and scissors were more similar to that of hands than other typical object categories, in the hand areas.

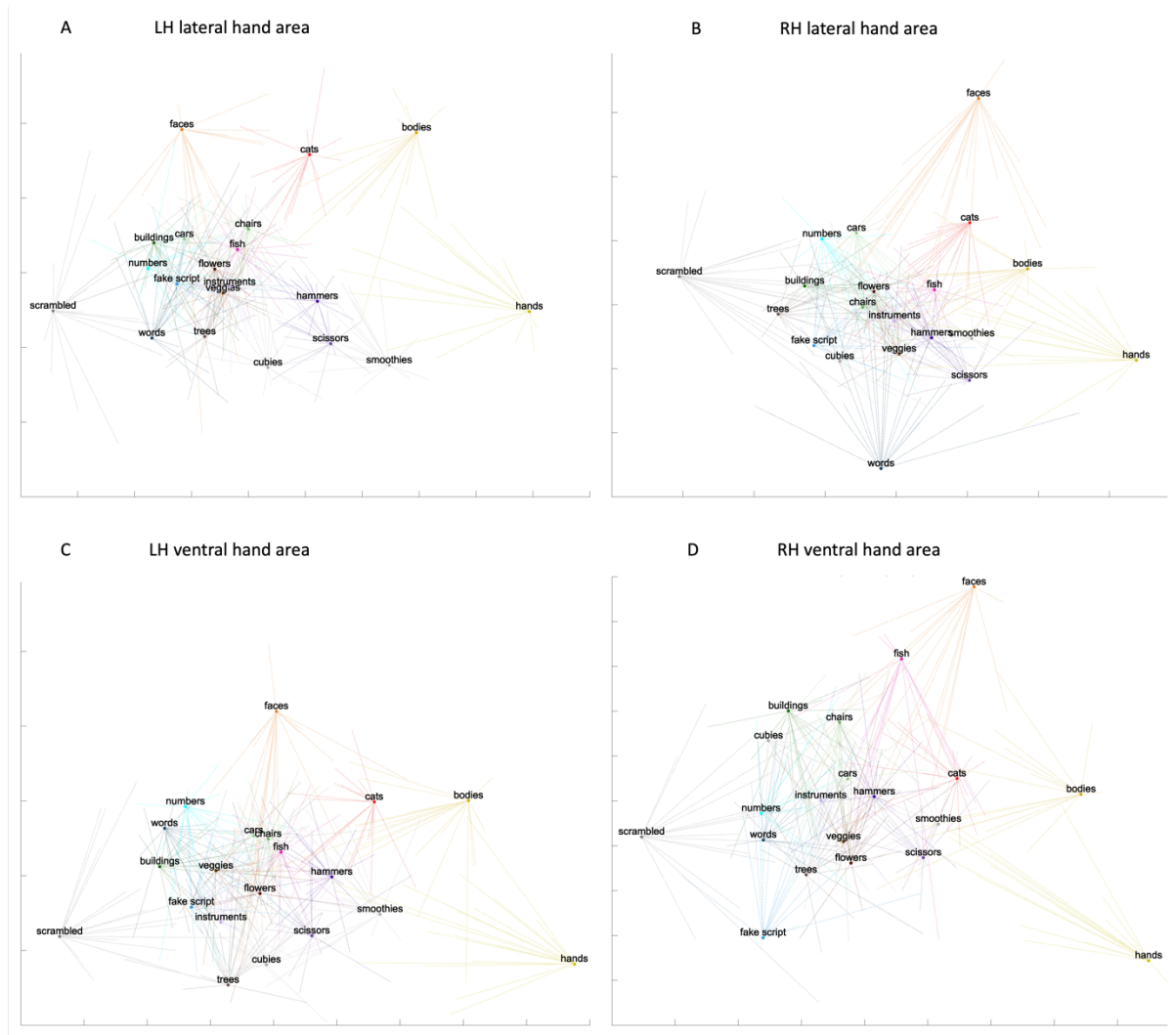


Figure 11. Visualization of the average multi-voxel pattern analysis results in a two-dimensional space using MDS, for the (A) left lateral hand area, (B) right lateral hand area, (C) left ventral hand area, and the (D) right ventral hand area. Each participant's MDS results were Procrustes transformed to the average MDS results and were shown by one line per category per subject in each of the four spaces.

Second, we investigated the results from the tool areas (see Figure 12). We could define a left lateral tool area in 18 subjects (mean size: 65.25 voxels, standard deviation: 37.7), a left ventral in 18 subjects (mean size: 18.42 voxels, standard deviation: 14.78), a right lateral in all 19 subjects (mean size: 82.16 voxels, standard deviation: 140.95) and a right ventral in 15 subjects (mean size: 13.37 voxels, standard deviation: 25.11).

We tested if the distance between tools and hands was significantly (i.e., $p < .01$ based on Bonferroni correction) smaller than between tools (including hammers and scissors) and objects (including chairs, cars and musical instruments). This was not the case (left lateral tool area: $t(17)=-1.75$, $p=.1$; left ventral: $t(17)=-0.47$, $p=.65$; right lateral: $t(18)=1.05$, $p=.31$; right ventral: $t(14)=-0.79$,

$p=.45$). Thus, the neural representations of hands and of typical object categories were equally similar to the neural representation of tools in the tool areas.

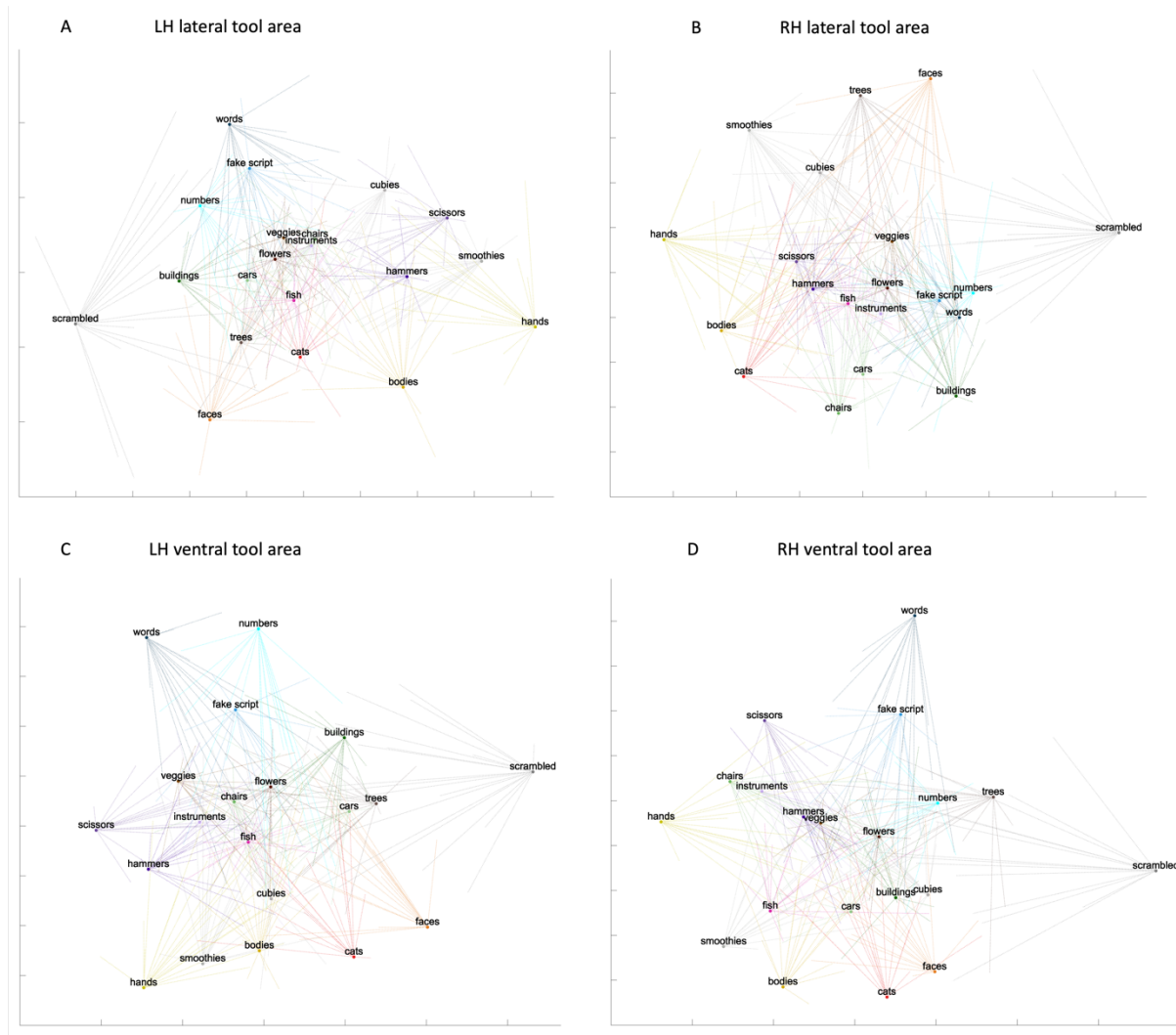


Figure 12. Visualization of the average multi-voxel pattern analysis results in a two-dimensional space using MDS, for the (A) left lateral tool area, (B) right lateral tool area, (C) left ventral tool area, and the (D) right ventral tool area. Each participant's MDS results were Procrustes transformed to the average MDS results and were shown by one line per category per subject in each of the four spaces.

Discussion

Research on the hand and tool areas and the overlap between them in OTC has been limited to the use of a standard spatial resolution with 3T. In this study, we reinvestigated these areas and their overlap at a high spatial resolution using 7T fMRI in 19 participants. Previous studies also applied several spatial preprocessing steps that might interfere with results. By using 7T fMRI, we were able to omit these preprocessing steps (spatial normalization and smoothing). Additionally, Bracci and colleagues (2016) defined hand and tool areas versus scrambled objects. This was a much less strict contrast than in their previous studies (where they compared versus chairs or versus animals and even versus bodies). This could be a reason as to why they could in 2016 pick up on left/right ventral hand and tool areas, and specifically could also include a right lateral tool area (which they excluded in 2012). In contrast, in our study, we used a strict contrast (hands or tools versus all other categories of the experiment, including chairs, bodies, cats, etcetera) and could locate all these areas as well in each participant. This result suggested that with our dataset, we could re-examine the hand and tool selectivity and their overlap, avoiding several of the drawbacks of existing studies. We explored their response profile to several categories that were included in the experiment. As a comparison, we also identified body, face and object (chair) areas. After, we examined the overlap between these areas in all parts of OTC. Finally, we constructed the representational space (using all categories included in the experiment) of the hand and tool areas using multi-voxel pattern analysis. In this space, we again used the body, face and object categories as a comparison.

The hand areas in all four parts of OTC were identified in all participants (except in right ventral OTC in 15 subjects). The finding of hand areas in all parts of OTC is a first exciting finding. Another study by Bracci and colleagues (2016) demonstrated the expansion of the left lateral hand area to the right hemisphere and to the ventral OTC, after failing to pick up on these areas in previous studies. These four hands areas all differentiated between tools and other typical objects: the response to tools was higher than to objects. This coincides with results from the left lateral hand area from Bracci and colleagues (2010). In their 2012 study two, they also included tools and objects, but they did not explicitly report if the left and right lateral hand area that they identified, differentiated tools from objects. However, from the bar graph of their results, it would seem so in the left lateral hand area, but not in the right lateral hand area. The latter is in contrast with our results. In their study in 2016, Bracci and colleagues found similar results as in their previous studies: a higher response to tools than to other objects in the left lateral hand area, but not in any of the other three hand areas. This is in contrast with our results, where all hand areas (not only the left lateral) made a differentiation between tools and other objects. It is possible that this effect is less clear in these other three hand areas (left ventral and right lateral and ventral OTC) and therefore, the previous studies could not pick

up on it, but due the advantages that 7T fMRI offers, we did. This finding suggests that tools are an important category to the hand areas in all parts of OTC, worth differentiating from other types of objects.

The hand areas also responded more to bodies than to tools (except in left ventral OTC). In the study by Bracci and colleagues (2010), this effect was not found in the left lateral hand area. They defined this hand area using a strict contrast (hands versus bodies, body parts and chairs; somewhat similar to the contrast we used except we contrasted against even more categories). In that study, using such a strict contrast, they only found the left lateral hand area and could not locate the three other hand areas, which we could. The lack of the effect (higher response to bodies than tools) in their study could thus be due to a lack in power. In Bracci and colleagues (2012), study one (contrast: hands versus animals), they found that in the left lateral hand area (but not in the right lateral hand area) tools was the second preferred category, after hands and before animals. This finding is possibly in contrast to our results. Then again, animals are not the same as human bodies so making the comparison between our finding and theirs is not possible. In study two (contrast: hands versus chairs) of Bracci and colleagues (2012), a test between tools and bodies/body parts was not specifically mentioned, but their bar graph suggests that in the left lateral hand area bodies and/or body parts are either preferred more than tools (and are thus the second preferred category) or are equally preferred to tools. This would correspond with our results. In the right lateral hand area, they found bodies to be preferred more than tools, again in accordance with our results. In 2016, Bracci and colleagues then found that bodies scored higher than tools, in all hand areas. These recent findings in combination with ours thus show that in the left/right lateral and right ventral hand areas, bodies is preferred above tools. It is inconclusive however if this is also true in the left ventral hand area, as we failed to confirm the effect here. These results are not unexpected: hands and bodies are both animate categories and are linked in that hands are a body part. Therefore, it makes sense that hand areas would also respond highly to bodies.

In the lateral hand areas, the response to tools was also higher than to faces. Faces are an animate category, just like hands and bodies, but nonetheless in the lateral hand areas, tools, that are inanimate, were preferred above faces (but not above bodies). In the ventral hand areas, there was no difference in response to tools and faces, which is also surprising given that one would expect the hand areas to prefer animate categories more than inanimate categories and not equally. We cannot compare this result to results from Bracci and colleagues (2010; 2012; 2016) as they did not include faces.

Tool areas were identified in most participants, except in right ventral OTC where only 8 subjects showed a tool area. In their 2012 study, Bracci and colleagues had trouble defining the right lateral

tool area and only defined the left one. In their 2016 study however, they could define right tool areas in most of their participants, as much as they could define right hand areas. The reason for this lack of a tool area in right ventral OTC in more than half of our participants remains unclear. One would expect, based on animacy, that these tool areas would prefer objects more than hands, bodies or faces, which are all animate categories. Overall, these areas actually preferred hands specifically to all other categories of interest (faces, bodies, objects). The effect that hands were preferred above all other categories, was less pronounced in the right hemisphere tool areas, where hands were not preferred more than bodies, but still more than faces and objects (except in the right ventral tool area). These findings are more specific but overall consistent with what was reported by Bracci and colleagues (2012), study one: they found that the left lateral tool area (they did not analyze other tool areas) responded more to hands than to animals or scenes. In study two, they also found hands to be the most preferred category, also more than bodies and body parts. A more recent study by Bracci and colleagues in 2016, investigated all tool areas. They found the response to hands to be higher than to bodies and objects (they did not include faces) in the left lateral tool area, just like in their previous studies and our study. In the right lateral tool area, they found that the response to bodies was actually higher than hands, but hands still scored higher than objects. This is similar to our results, as hands did not score higher than bodies, but still higher than objects. In the ventral left tool area, they found no higher response to hands than to bodies or objects. This is in contrast to our results because hands did score higher than objects and bodies. In the ventral right tool area, they found that bodies and objects scored equally higher than hands. We found that hands did not score higher than bodies or objects (only higher than faces). However, it should be taken into account that this area could not be identified in more than half of the subjects, which may have impacted the statistics. Overall, their and our findings suggest that only the left hemisphere tool areas needed to differentiate between hands and bodies. Importantly, to all tool areas, hands were special in that they always received a higher response compared to some other categories (e.g., faces and objects) and thus were differentiated from these other categories.

Face, body and object areas were also identified as a control. Most face and body areas could not, unlike the hand areas, differentiate between tools and other objects (exceptions were the right ventral face area and right lateral body area). Bracci and colleagues (2010) also confirmed that body-selective voxels did not differentiate between tools and objects (chairs). In Bracci and colleagues (2012) study two, they also identified lateral body areas, but they did not mention testing if these areas differentiated between tools and objects (chairs). From their figure, it seems that the left lateral body area might have been able to (in contrast with our results), but the right lateral body area did not (also in contrast with our results as this was the only body area that we found that could make the

differentiation). They did not include face areas for comparison (Bracci et al., 2010; Bracci et al., 2012; Bracci et al., 2016). The finding here that the face areas also overall, like the body areas, do not differentiate between tools and other objects, suggests that the hand areas are special in this regard. Only to these hand areas it is important to differentiate between tools and other objects. The identified object (chair) areas overall preferred other typical objects more than it did tools. Although in study one by Bracci and colleagues (2012), a LO region was defined, they did not conduct a test similar to ours so we cannot compare results.

Then, overlap between areas was analyzed. The overlap between hands and tools was specific in the sense that this overlap was larger than the overlap between faces and tools, between bodies and tools, and between hands and objects, in the left hemisphere. Bracci and colleagues in study two (2012) found in left lateral OTC that the overlap between hands and tools was larger than tools and bodies, corresponding to our results. In the right lateral OTC, we found that the overlap was somewhat specific: it was larger than the overlap between faces and tools and between hands and objects, but not larger than the overlap between bodies and tools. In right ventral OTC, the overlap was not specific. However, the lack of participants that could be used for tests in this area, due to a lack of a tool area here, should be taken into account, as this could have impacted the statistics. Finally, we found that the overlap between hands and tools was not larger in lateral versus ventral OTC, not in the left or the right hemisphere. We did find a difference between hemispheres: there was more overlap between hands and tools in the left lateral OTC as compared to the right lateral OTC. The same could not be said for the left versus right ventral OTC, but again the lack of participants for this test (due to missing a right ventral tool area) should be taken into account.

At last, we investigated the representational space of the left/right lateral/ventral hand and tool areas. In the hand areas, we tested the similarity between neural representations of tools and hands in comparison to the similarity between hands and typical objects. We found that the representation of tools was more similar than that of objects to the representation of hands. In the tool areas, we found that the neural representations to hands and to typical objects were all equally similar to the representation of tools. We may have expected tools and hands to have a similar neural representation based on previous studies, where there was no difference in response to hands and tools or even a higher response to hands than to tools in a univariate analysis of the tool areas (Bracci et al., 2012; Bracci et al., 2016). Bracci and colleagues (2012, 2016) also performed MVPA, but either only in a general lateral OTC and not in hand and tool areas separately, or not using methods/tests like ours, therefore we cannot compare results.

It's important to note that, especially in the univariate analyses, some category-selective areas could not be defined in many of the participants (e.g., the right ventral tool area) which influenced the

statistics that were carried out in these areas (the tests on its response profile) or that related to overlap between this and other areas.

A main benefit of this study was the use of 7T fMRI. This allowed the use of higher spatial resolution (van der Zwaag et al., 2016), which is important especially for the finding of overlap between hand and tool areas, as this might also be a result of a lack in spatial resolution (Bracci et al., 2012). 7T fMRI also provided high quality data (improved signal-to-noise ratio, van der Zwaag et al., 2016). We could thus omit spatial normalization and smoothing of the data. This again was critical for the analysis of overlap between hand and tool areas, as smoothing may lead to averaging areas together that actually lie apart from each other on the surface (Weiner & Grill-Spector, 2013). Combined with the higher specificity of how we defined hand and tool selectivity (against many other categories, not against scrambled like Bracci et al., 2016), these characteristics of 7T data and analyses could all have been reasons to not be able to confirm and even strengthen the evidence from prior studies about hand and tool selectivity and their overlap.

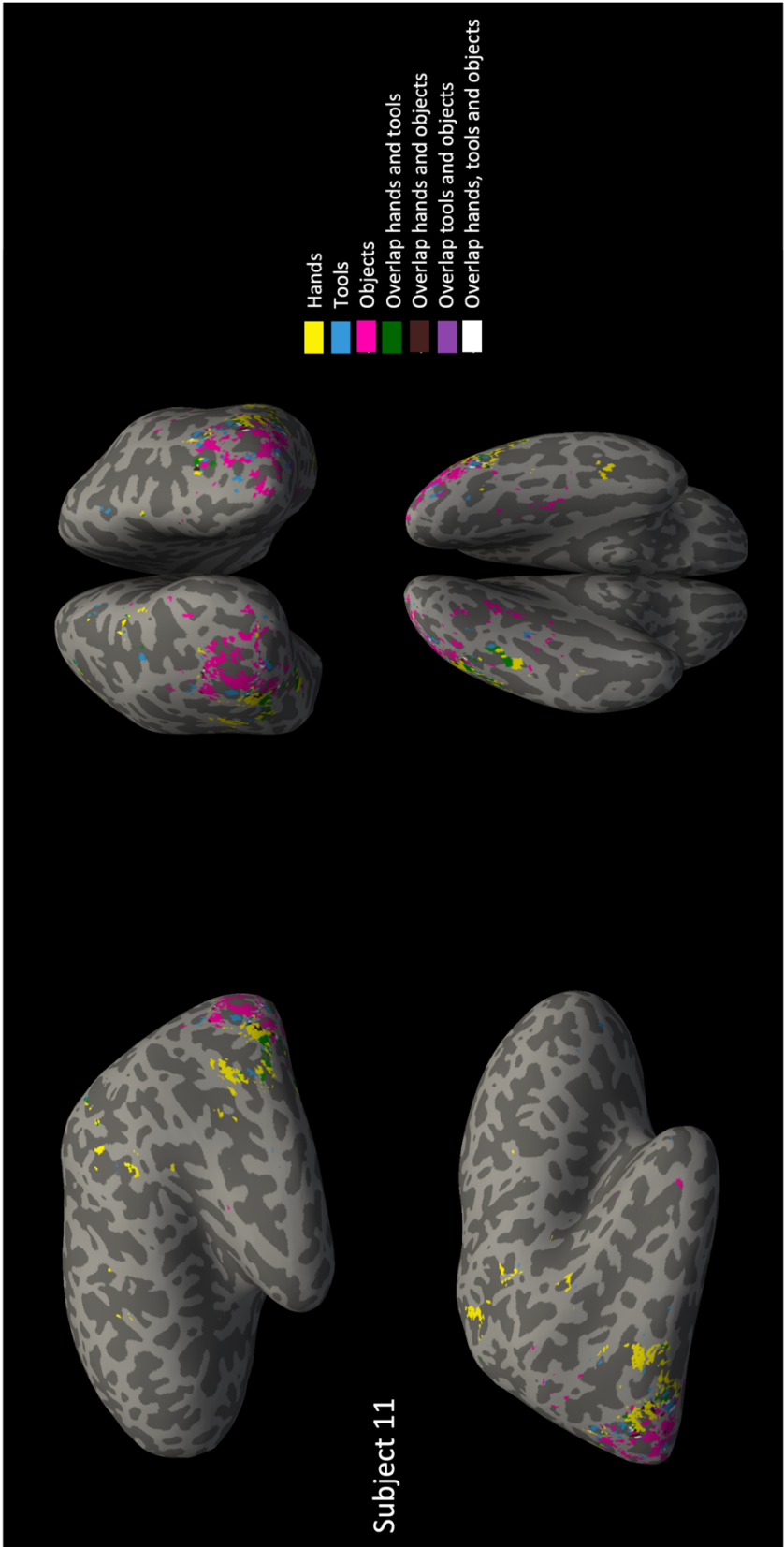
Another benefit of this study was that, in the creation of our stimuli, we ensured that bodies would not contain hands, as they do not contain faces either since these categories are supposed to be separate. We do not suggest that this detail was of influence in previous studies, but we do believe this experimental detail would be good practice for future studies since the influence of it may not be obvious.

This work has added to previous findings in several ways. It used a higher spatial resolution than previous studies and omitted spatial preprocessing steps that could influence the results and maybe even make all overlap between hand and tool selectivity disappear. To the contrary, our work showed that not only the left lateral hand area, but all hand areas differentiated between tools and other objects. In addition, we compared the tool responses not only with hand responses but also with several other human categories, including bodies and faces. We confirmed that bodies were preferred more than tools by the hand areas. In contrast, we showed for the first time that hand areas responded more to tools than to faces. This study also demonstrated that not only the left lateral tool area, but overall, all tool areas differentiated between hands, faces and other types of objects than tools. The left hemisphere tool areas also differentiated between hands and bodies. Our study is also the first to our knowledge to investigate the overlap (its specificity and size) between hand and tool areas beyond the left lateral OTC (as done by Bracci et al., 2012). It showed that the overlap was specific to these two categories (hands and tools) in the left hemisphere, and less so in the right hemisphere. In addition, it showed that the size of the overlap was not significantly different between lateral and ventral OTC (per hemisphere), but the overlap was larger in the left versus right lateral OTC (not in left versus right ventral OTC). Finally, we also constructed the representational space of all hand and tool

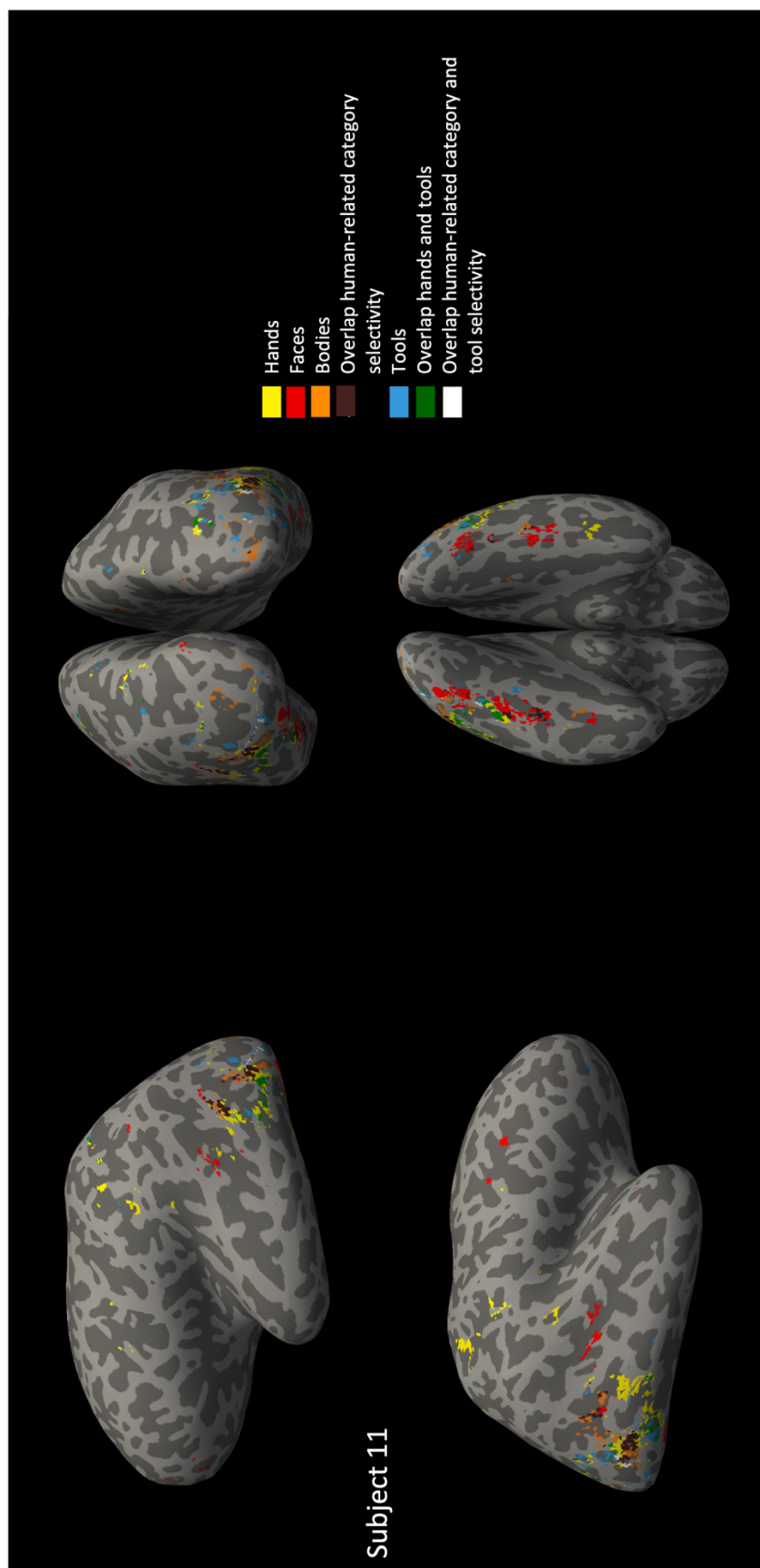
areas and compared the distance between neural representations of several category pairs. This was not done in previous studies. We found that in the hand areas, the representation for tools was more similar than the representation for objects, to the representation of hands. In the tool areas, the representation for hands and for objects was equally similar to that of tools.

Future work could investigate further these different hand and tool areas to understand what specific roles they play and how these roles differ from each other and why. Although we used a higher spatial resolution than previous studies and demonstrated overlap between hand and tool areas, it is still possible that an even higher spatial resolution would show these areas are in fact separate and just lie really close, or that there is no overlap at the level of single-neuron selectivity. Also, functional connectivity studies between these four hand and tool areas with each other and other regions (e.g., action-related regions) would be a useful addition to the field to further the understanding of the roles of these areas in OTC and beyond.

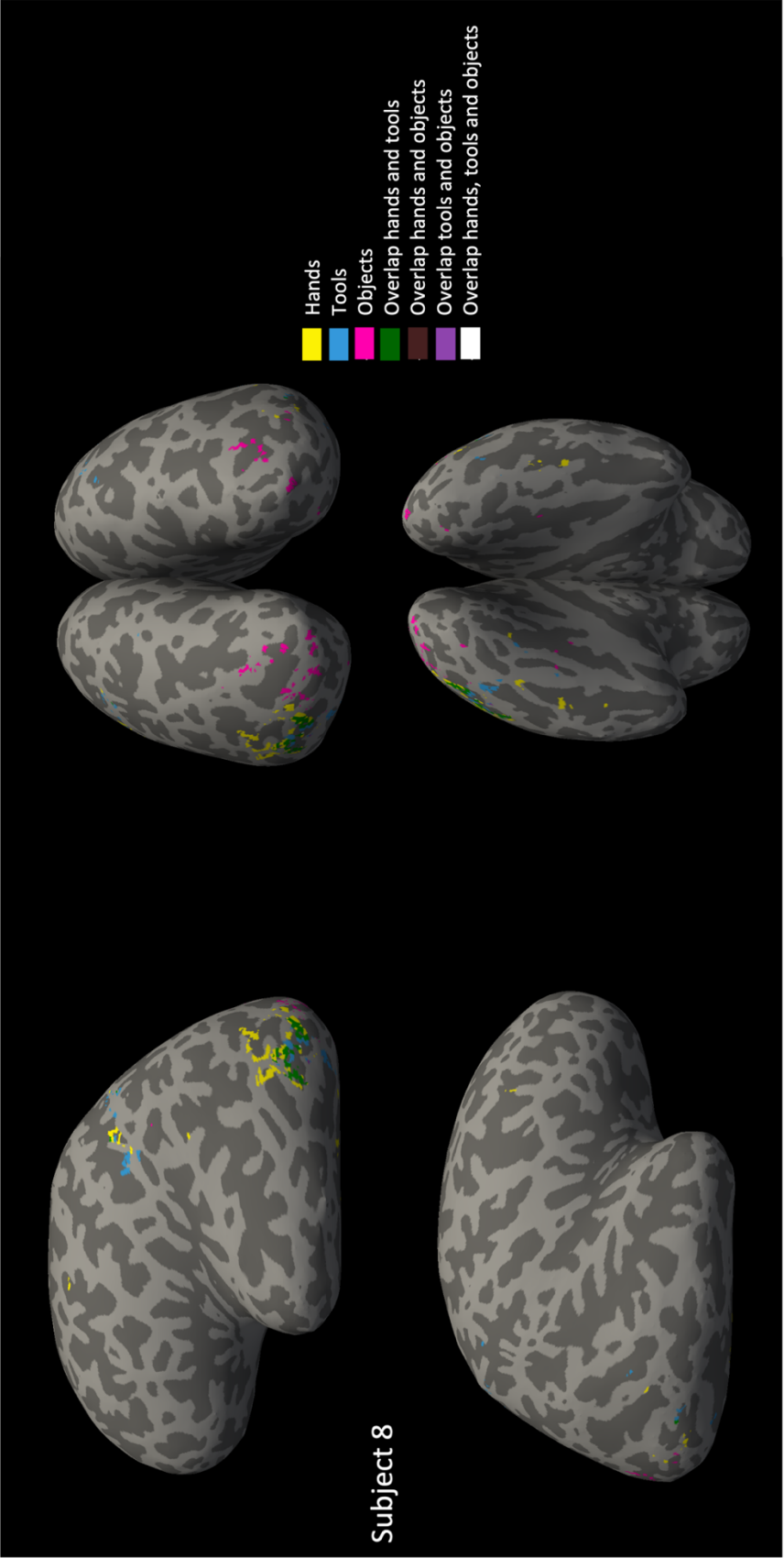
Supplementary material



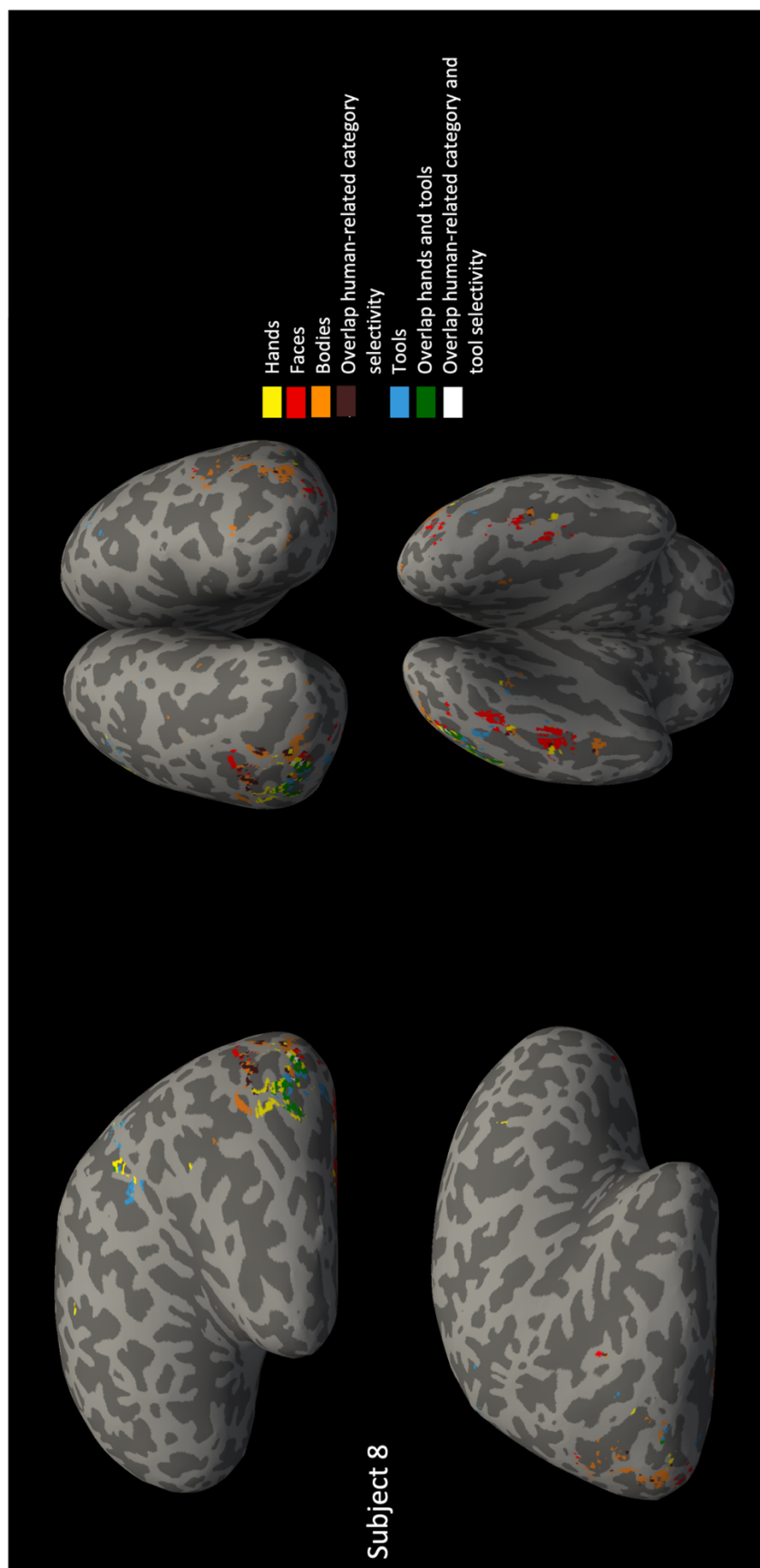
Supplementary Figure 1. Hand, tool and object areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, tools in blue, objects (chairs) in pink, overlap between hands and tools in green, between hands and objects in brown, between tools and objects in purple and between hands, tools and objects in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 11 (color legend also on the right of the figure).



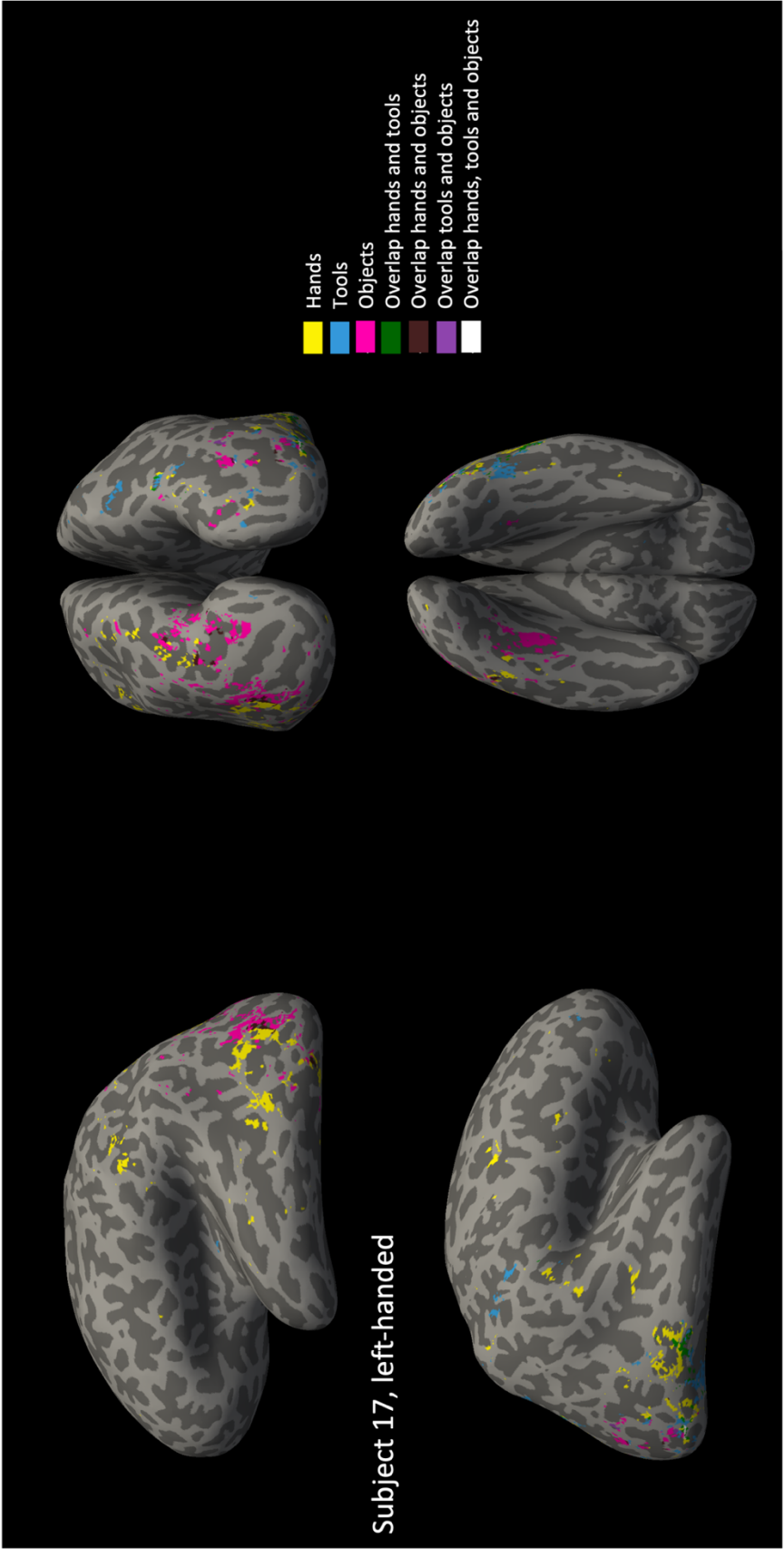
Supplementary Figure 2. Hand, face, body and tool areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, faces in red, bodies in orange, overlap between some or all of these three animate categories in brown, tools in blue, overlap between hands and tools in green, and between animate categories and tools in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 11 (color legend also on the right of the figure).



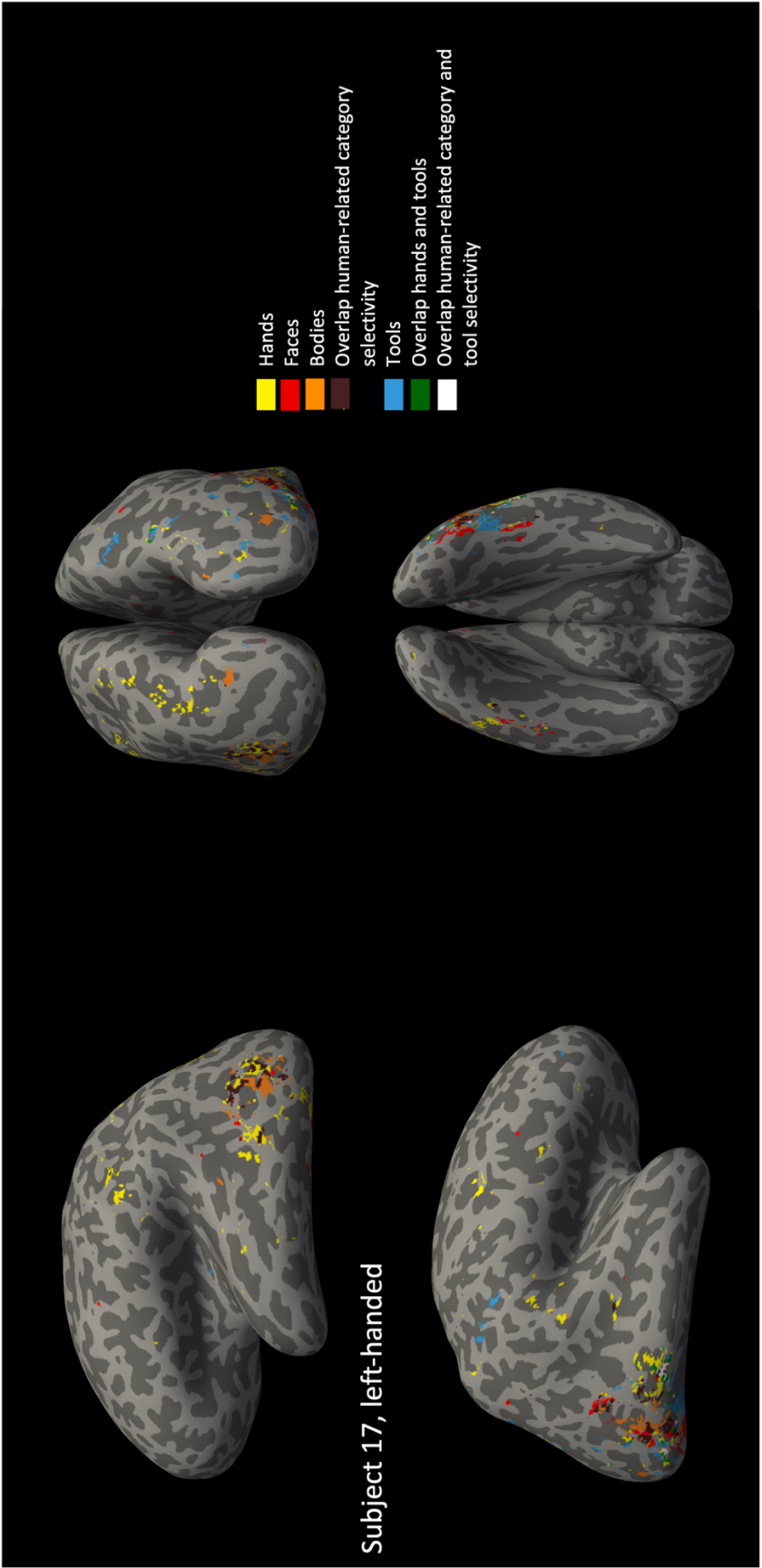
Supplementary Figure 3. Hand, tool and object areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, tools in blue, objects (chairs) in pink, overlap between hands and tools in green, between hands and objects in brown, between tools and objects in purple and between hands, tools and objects in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 8 (color legend also on the right of the figure).



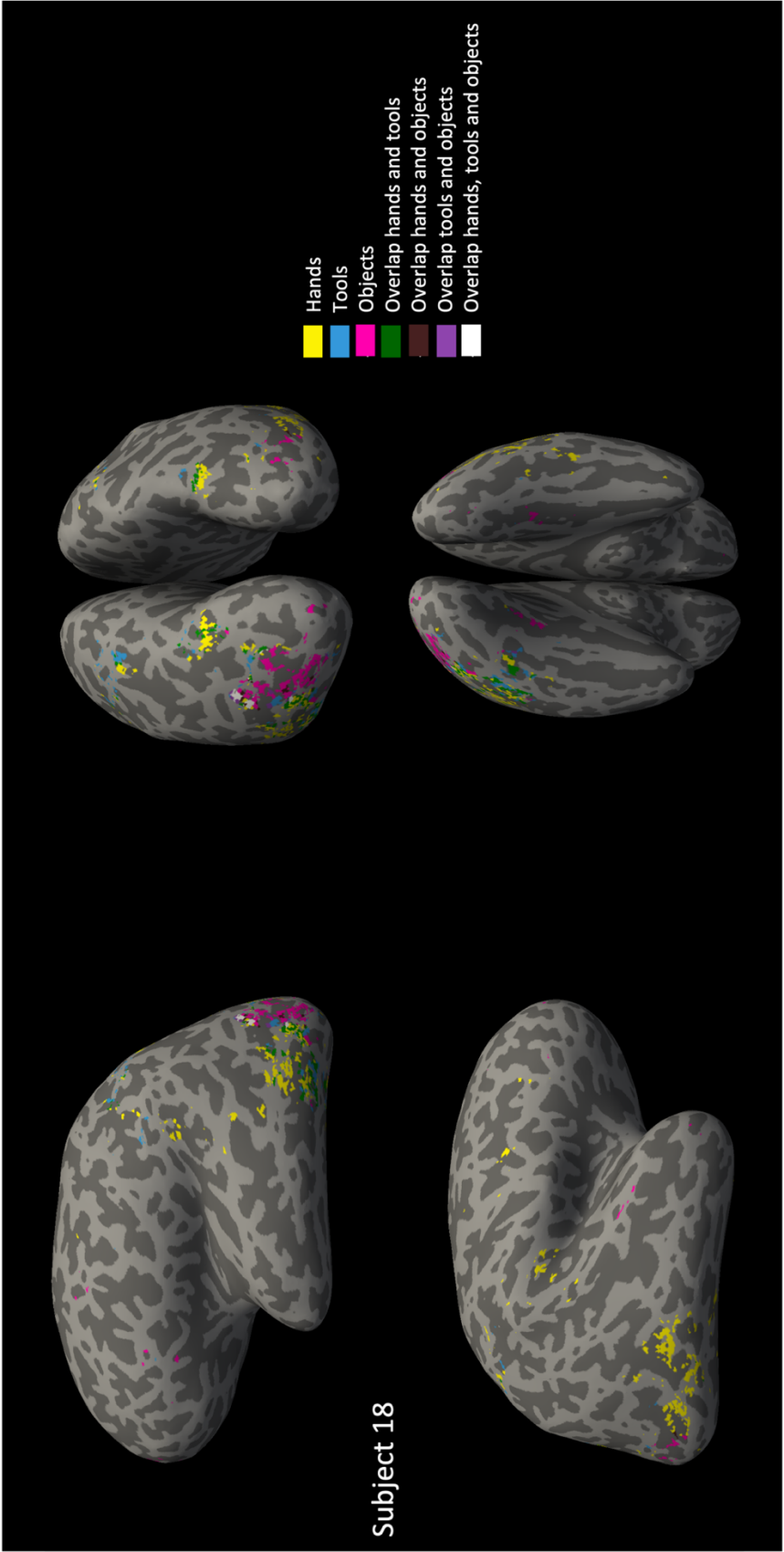
Supplementary Figure 4. Hand, face, body and tool areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, faces in red, bodies in orange, overlap between some or all of these three animate categories in brown, tools in blue, overlap between hands and tools in green, and between animate categories and tools in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 8 (color legend also on the right of the figure).



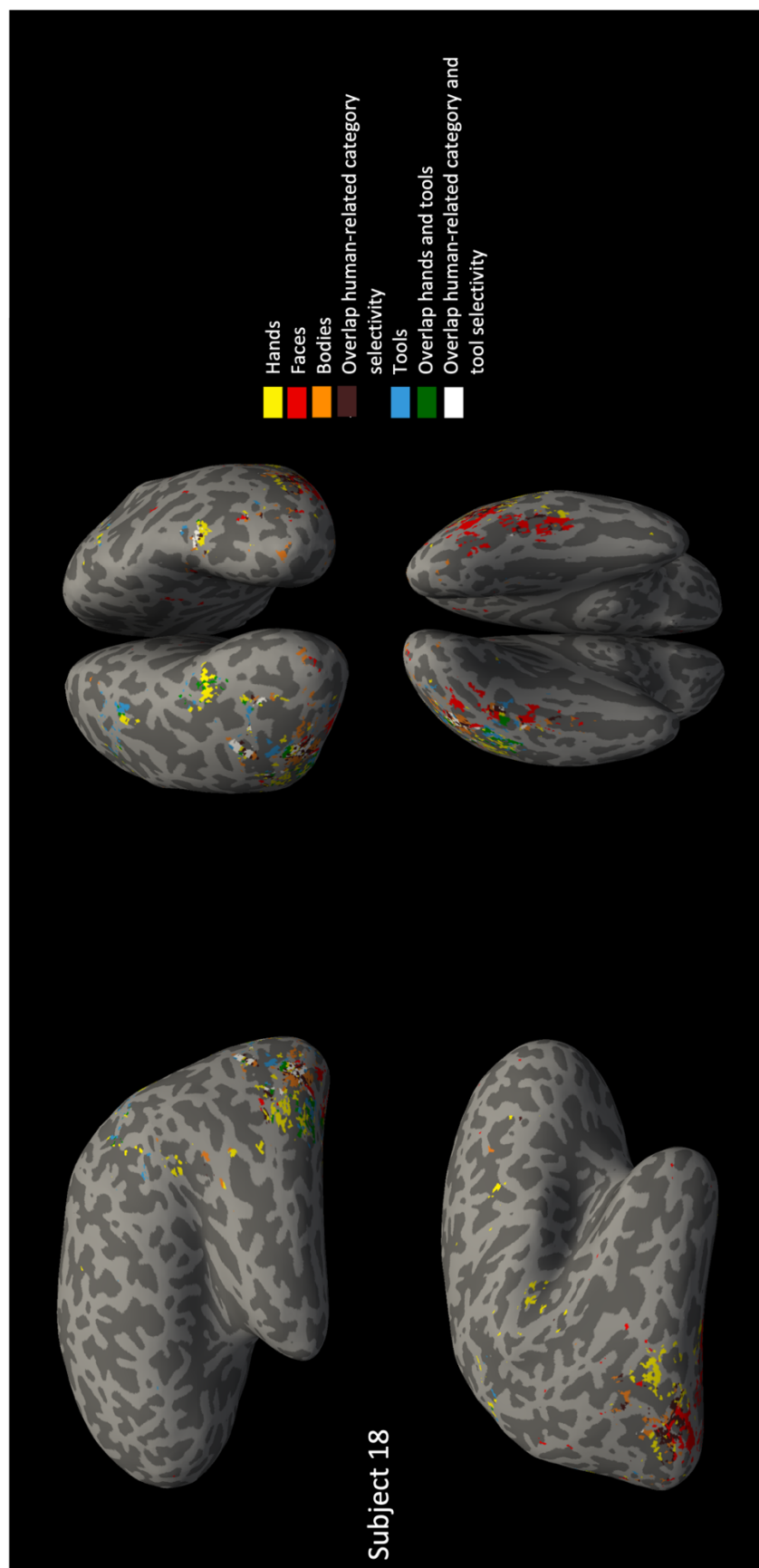
Supplementary Figure 5. Hand, tool and object areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, tools in blue, objects (chairs) in pink, overlap between hands and tools in green, between hands and objects in brown, between tools and objects in purple and between hands, tools and objects in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 17 (color legend also on the right of the figure).



Supplementary Figure 6. Hand, face, body and tool areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, faces in red, bodies in orange, overlap between some or all of these three animate categories in brown, tools in blue, overlap between hands and tools in green, and between animate categories and tools in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 17 (color legend also on the right of the figure).



Supplementary Figure 7. Hand, tool and object areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, tools in blue, objects (chairs) in pink, overlap between hands and tools in green, between hands and objects in brown, between tools and objects in purple and between hands, tools and objects in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 18 (color legend also on the right of the figure).



Supplementary Figure 8. Hand, face, body and tool areas and their overlap, determined by a contrast of one versus all other categories except fixation, $p < 0.05$, FWE corrected (hands in yellow, faces in red, bodies in orange, overlap between some or all of these three animate categories in brown, tools in blue, overlap between hands and tools in green, and between animate categories and tools in white), shown upon annotated left lateral (top left), posterior (top right), right lateral (bottom left) and ventral (bottom right) brain surface of participant 18 (color legend also on the right of the figure).

References

- Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, 8. <https://doi.org/10.3389/fninf.2014.00014>
- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to “building” stimuli. *Neuron*, 21(2), 373–383. [https://doi.org/10.1016/s0896-6273\(00\)80546-2](https://doi.org/10.1016/s0896-6273(00)80546-2)
- Avants, B., Epstein, C., Grossman, M., & Gee, J. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1), 26–41. <https://doi.org/10.1016/j.media.2007.06.004>
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(21), 9087–9092. <https://doi.org/10.1073/pnas.0703300104>
- Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Bracci, S., Cavina-Pratesi, C., Connolly, J. D., & Ietswaart, M. (2016). Representational content of occipitotemporal and parietal tool areas. *Neuropsychologia*, 84, 81–88. <https://doi.org/10.1016/j.neuropsychologia.2015.09.001>
- 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. <https://doi.org/10.1152/jn.00619.2011>
- 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. <https://doi.org/10.1152/jn.00215.2010>
- Bracci, S., & Peelen, M. V. (2013). Body and object effectors: the organization of object representations in high-level visual cortex reflects body-object interactions. *The Journal of Neuroscience*, 33(46), 18247–18258. <https://doi.org/10.1523/JNEUROSCI.1322-13.2013>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436. <https://doi.org/10.1163/156856897x00357>
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS One*, 5(5), e10773. <https://doi.org/10.1371/journal.pone.0010773>

- Brodoehl, S., Gaser, C., Dahnke, R., Witte, O. W., & Klingner, C. M. (2020). Surface-based analysis increases the specificity of cortical activation patterns and connectivity results. *Scientific Reports*, 10(1), 5737. <https://doi.org/10.1038/s41598-020-62832-z>
- Caffarra, S., Karipidis, I. I., Yablonski, M., & Yeatman, J. D. (2021). Anatomy and physiology of word-selective visual cortex: from visual features to lexical processing. *Brain Structure & Function*, 226(9), 3051–3065. <https://doi.org/10.1007/s00429-021-02384-8>
- 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. <https://doi.org/10.1038/13217>
- Coggan, D. D., Liu, W., Baker, D. H., & Andrews, T. J. (2016). Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. *NeuroImage*, 135, 107–114. <https://doi.org/10.1016/j.neuroimage.2016.04.060>
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., & Michel, F. (2000). The visual word form area. *Brain*, 123(2), 291–307. <https://doi.org/10.1093/brain/123.2.291>
- Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, 125(Pt 5), 1054–1069. <https://doi.org/10.1093/brain/awf094>
- Cohen, M. A., Alvarez, G. A., Nakayama, K., & Konkle, T. (2017). Visual search for object categories is predicted by the representational architecture of high-level visual cortex. *Journal of Neurophysiology*, 117(1), 388–402. <https://doi.org/10.1152/jn.00569.2016>
- Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, 10(4–5), 171–178. [https://doi.org/10.1002/\(sici\)1099-1492\(199706/08\)10:4/5<171::aid-nbm453>3.0.co;2-l](https://doi.org/10.1002/(sici)1099-1492(199706/08)10:4/5<171::aid-nbm453>3.0.co;2-l)
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2), 179–194. <https://doi.org/10.1006/nimg.1998.0395>
- 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. <https://doi.org/10.1126/science.1063414>
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, 392(6676), 598–601. <https://doi.org/10.1038/33402>
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R.

- A., & Gorgolewski, K. J. (2018). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1), 111–116. <https://doi.org/10.1038/s41592-018-0235-4>
- Esteban, O., Markiewicz, C. J., Goncalves, M., Provins, C., Kent, J. D., DuPre, E., Salo, T., Ciric, R., Pinsard, B., Blair, R. W., Poldrack, R. A., & Gorgolewski, K. J. (2018). fMRIPrep. *Software*. Zenodo. <https://doi.org/10.5281/ZENODO.852659>
- Fonov, V. S., Evans, A. C., McKinstry, R. C., Alml, C. R., & Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47, S102. [https://doi.org/10.1016/s1053-8119\(09\)70884-5](https://doi.org/10.1016/s1053-8119(09)70884-5)
- Garcea, F. E., & Mahon, B. Z. (2014). Parcellation of left parietal tool representations by functional connectivity. *Neuropsychologia*, 60, 131–143. <https://doi.org/10.1016/j.neuropsychologia.2014.05.018>
- Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine*, 44(1), 162–167. [https://doi.org/10.1002/1522-2594\(200007\)44:1<162::aid-mrm23>3.0.co;2-e](https://doi.org/10.1002/1522-2594(200007)44:1<162::aid-mrm23>3.0.co;2-e)
- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, 5, 13. <https://doi.org/10.3389/fninf.2011.00013>
- Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., ... Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3(1). <https://doi.org/10.1038/sdata.2016.44>
- Gorgolewski, K. J., Esteban, O., Markiewicz, C. J., Ziegler, E., Ellis, D. G., Notter, M. P., Jarecka, D., ..., & Ghosh, S. (2018). Nipype. *Software*. Zenodo. <https://doi.org/10.5281/zenodo.596855>
- Gras, V., Vignaud, A., Amadon, A., Le Bihan, D., & Boulant, N. (2017). Universal pulses: A new concept for calibration-free parallel transmission. *Magnetic Resonance in Medicine*, 77(2), 635–643. <https://doi.org/10.1002/mrm.26148>
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1), 63–72. <https://doi.org/10.1016/j.neuroimage.2009.06.060>
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, 15(8), 536–548. <https://doi.org/10.1038/nrn3747>

- Harvey, A. K., Pattinson, K. T. S., Brooks, J. C. W., Mayhew, S. D., Jenkinson, M., & Wise, R. G. (2008). Brainstem functional magnetic resonance imaging: Disentangling signal from physiological noise. *Journal of Magnetic Resonance Imaging*, 28(6), 1337–1344.
<https://doi.org/10.1002/jmri.21623>
- Hutton, C., Josephs, O., Stadler, J., Featherstone, E., Reid, A., Speck, O., Bernarding, J., & Weiskopf, N. (2011). The impact of physiological noise correction on fMRI at 7 T. *NeuroImage*, 57(1), 101–112. <https://doi.org/10.1016/j.neuroimage.2011.04.018>
- 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. <https://doi.org/10.1006/nimg.2002.1132>
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17(11), 4302–4311. <https://doi.org/10.1523/jneurosci.17-11-04302.1997>
- Kasper, L., Bollmann, S., Diaconescu, A. O., Hutton, C., Heinzle, J., Iglesias, S., Hauser, T. U., Sebold, M., Manjaly, Z.-M., Pruessmann, K. P., & Stephan, K. E. (2017). The PhysIO toolbox for modeling physiological noise in fMRI data. *Journal of Neuroscience Methods*, 276, 56–72.
<https://doi.org/10.1016/j.jneumeth.2016.10.019>
- Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLoS Computational Biology*, 13(2), e1005350. <https://doi.org/10.1371/journal.pcbi.1005350>
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1–16.
- Knights, E., Mansfield, C., Tonin, D., Saada, J., Smith, F. W., & Rossit, S. (2021). Hand-selective visual regions represent how to grasp 3D tools: Brain decoding during real actions. *The Journal of Neuroscience*, 41(24), 5263–5273. <https://doi.org/10.1523/JNEUROSCI.0083-21.2021>
- Lanczos, C. (1964). Evaluation of noisy data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 1(1), 76–85. <https://doi.org/10.1137/0701007>
- Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods*, 264, 47–56.
<https://doi.org/10.1016/j.jneumeth.2016.03.001>
- Mahon, B. Z., Milleville, S. C., Negri, G. A. L., Rumiat, R. I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, 55(3), 507–520. <https://doi.org/10.1016/j.neuron.2007.07.011>

- 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. <https://doi.org/10.3389/fninf.2016.00027>
- Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *The Journal of Neuroscience, 26*(50), 13025–13036. <https://doi.org/10.1523/JNEUROSCI.2481-06.2006>
- 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. https://doi.org/10.1162/jocn_a_00411
- Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology, 93*(1), 603–608. <https://doi.org/10.1152/jn.00513.2004>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437–442. <https://doi.org/10.1163/156856897x00366>
- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage, 84*, 320–341. <https://doi.org/10.1016/j.neuroimage.2013.08.048>
- Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., Eickhoff, S. B., Hakonarson, H., Gur, R. C., Gur, R. E., & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage, 64*, 240–256. <https://doi.org/10.1016/j.neuroimage.2012.08.052>
- Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *The Journal of Neuroscience, 25*(47), 11055–11059. <https://doi.org/10.1523/JNEUROSCI.2621-05.2005>
- Striem-Amit, E., Vannuscorps, G., & Caramazza, A. (2017). Sensorimotor-independent development of hands and tools selectivity in the visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 114*(18), 4787–4792. <https://doi.org/10.1073/pnas.1620289114>
- Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*(5761), 670–674. <https://doi.org/10.1126/science.1119983>
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging, 29*(6), 1310–1320. <https://doi.org/10.1109/tmi.2010.2046908>

- Tustison, N. J., Cook, P. A., Klein, A., Song, G., Das, S. R., Duda, J. T., Kandel, B. M., van Strien, N., Stone, J. R., Gee, J. C., & Avants, B. B. (2014). Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *NeuroImage*, 99, 166–179.
<https://doi.org/10.1016/j.neuroimage.2014.05.044>
- Valyear, K. F., Cavina-Pratesi, C., Stiglick, A. J., & Culham, J. C. (2007). Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *NeuroImage*, 36 Suppl 2, T94–T108.
<https://doi.org/10.1016/j.neuroimage.2007.03.031>
- van der Zwaag, W., Schäfer, A., Marques, J. P., Turner, R., & Trampel, R. (2016). Recent applications of UHF-MRI in the study of human brain function and structure: a review: UHF MRI: Applications to Human Brain Function and Structure. *NMR in Biomedicine*, 29(9), 1274–1288.
<https://doi.org/10.1002/nbm.3275>
- Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, 137, 188–200.
<https://doi.org/10.1016/j.neuroimage.2015.12.012>
- 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. <https://doi.org/10.1007/s00426-011-0392-x>
- 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.
<https://doi.org/10.3758/BRM.42.3.671>
- Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., & Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 31(3), 1116–1128.
<https://doi.org/10.1016/j.neuroimage.2006.01.015>
- Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1), 45–57. <https://doi.org/10.1109/42.906424>

General Discussion

Objectives

This PhD aimed to investigate the neural basis of visual categorization with innovative methodologies and analyses. This PhD was built upon a rich history of research on the occipitotemporal cortex (OTC), that revealed several key findings about its functional organization. In short, OTC consists of category-selective areas and is organized according to several large-scale dimensions such as the animacy of categories.

The first objective of this PhD included comparing three popular analyses used on fMRI data, to better understand what these analyses can tell us about the data and allow future research to make an informed choice between them.

In this PhD, access to a 7T MRI scanner was provided, and we thus determined which research topics would benefit from the use of such a scanner and the benefits that it brings with it. The second objective included using the high spatial resolution of 7T fMRI, at the level of the individual subject, to investigate anatomically separate subareas of the visual word form area (VWFA) and how these areas relate to other category-selective areas.

The third objective of this PhD was to reveal if the hand and tool areas of OTC overlap at a higher spatial resolution obtained through 7T fMRI, and, with this higher signal-to-noise ratio, investigate the response profiles of these areas in the several parts of OTC.

To achieve the first objective, I first analyzed a dataset using the three popular analyses (univariate analysis, multivariate analysis, and functional connectivity analysis), and second, compared the results of these three analyses to each other.

To achieve the second and third objective, I designed a 7T fMRI study. This included running several pilot studies with both 3T and 7T fMRI, analyzing another 3T fMRI dataset that was already available in the lab (these steps are not presented in any further detail in this PhD thesis), and designing an experiment and appropriate stimuli. After data collection, the data was analyzed separately to achieve the second and third objective respectively. The choice of analyses was guided by the results of the study carried out in the context of the first objective.

Main findings

Comparing functional networks revealed by three different fMRI analyses

Using univariate and functional connectivity analysis, functional networks of the brain can be investigated. With the recent emergence of multivariate or multi-voxel pattern analysis and consequent representational similarity analysis, functional networks can now also be investigated in another manner (Kriegeskorte et al., 2008). To our knowledge, no study had compared the networks that could be built up with these three analyses, to understand if they reveal similar information about functional networks in the brain. To achieve this, we analyzed a dataset (see Lee Masson et al., 2018) using these three types of analyses and then compared the results by (partial) correlation and regression models. This study has been published (Pillet et al., 2019).

The study revealed that networks from each type of analysis are highly similar: regions of interest from the somatosensory-motor network were activated similarly, represented task information similarly and were connected to each other; as based on the univariate, multivariate and functional connectivity analyses respectively. Other networks (pain, social-cognitive, and visual networks) were also investigated, and comparable results were found. These findings support the idea that areas that represent task related information are also functionally connected to each other.

This finding was not surprising. These analyses share several features that can explain this result. For example, univariate and functional connectivity analysis both involve the averaging of neural signal across the voxels of a brain region. As another example, multivariate and functional connectivity analysis both involve a measure of similarity between brain regions. Also based on previous studies, that compare some of the three methods, this did not come as a surprise. Zeharia and colleagues (2015) demonstrated for example similar findings from multivariate and functional connectivity analyses.

However, our study also demonstrated that some of the information as uncovered by one method of analysis, could not be predicted by the other methods of analyses. Thus, each method can reveal unique information about functional networks in the brain. As an example, there were moderate correlations between social-cognitive brain areas and somatosensory-motor areas in both the networks revealed by univariate and multivariate analyses, but not in the network as obtained by functional connectivity analysis. As another example, the activation of the areas in the social-cognitive and visual network were similar, but the neural patterns of those same areas were not similar, nor were they functionally connected.

This finding could also be expected as these methods, although they all share some similarities, are still different. For example, although both multivariate and functional connectivity analyses correlate between areas, they correlate different things: the average neural signal across time versus

representational dissimilarity matrices. Also based on previous studies, comparing some of the three methods to one another, this finding was not surprising. For example, studies showed discrepancies in results from univariate versus multivariate results (Mur et al., 2009; Raizada et al., 2010). It is important to note that not all similarities or dissimilarities between the different methods of analyses relate to cognitive functioning. They may also come from other aspects such as vasculature.

The observation of unique info revealed by each type of analysis should be taken into account when studies make implications based on their results. More specifically, researchers should keep in mind how the methods differ. For example, the multivariate method reveals if brain areas represent similar task information whereas functional connectivity is focused on how brain areas together form a functional network to process the task. Even though there is a lot of similarity between them, when choosing methods of analyses, researchers should be mindful about how the methods are different with regards to what they can reveal about the functional organization of the brain, and how that relates to their research questions and goals. Only then can they choose appropriate methods.

The visual word form area

One of the category-selective areas in ventral OTC is the visual word form area (VWFA). The area appears in the posterior occipitotemporal sulcus (OTS) up to the middle of the fusiform gyrus (Yeatman et al., 2021). This region of word selectivity was recently divided into subareas: a posterior area (pOTS or VWFA-1) and an anterior area (mOTS or VWFA-2) (Lerma-Usabiaga et al., 2018; White et al., 2019). These subareas were not defined based on visible anatomical boundaries between them, but based on functional differences. Later prominent reviews then indicated the need for a study on the VWFA that uses a high spatial resolution and works at the level of the individual subject (without spatial normalization or smoothing), to determine if there is an anatomical basis for this functional division of the VWFA into subareas (Caffarra et al., 2021; Yeatman et al., 2021). To achieve that, we employed 7T fMRI to map word selectivity on the brains of 19 participants. We also mapped other types of category selectivity as there may be consistent relationships between word selectivity and other types, as seen for example for face and body areas (Weiner & Grill-Spector, 2010). For each of these areas, we also analyzed its response profile using a split-half analysis. On top of that, given that we included up to 20 categories that varied on several organizational dimensions like animacy, we could also investigate word selectivity in the representational space of OTC.

In the left hemisphere of OTC, we located three subareas of the VWFA: one posterior from lateral to medial ventral OTC, one in the middle and mostly on the lateral side, and one anterior and also rather on the lateral side. The second and third correspond to the VWFA-1 and 2 (Lerma-Usabiaga et al., 2018; White et al., 2019). The first subarea corresponds to a cluster of word selectivity that in the

literature was defined separately from the VWFA, because this area prefers letters but not specifically letter combinations that form words (James et al., 2005; Vinckier et al., 2007; Wong et al., 2009; Strother et al., 2015; Yeatman et al., 2021). We also identified the pFus word area, a cluster of word selectivity that was located in between or next to face area pFus/FFA-1, which to our knowledge has not been defined before as a subarea of the VWFA. This area may not differ functionally from either the VWFA-1 or 2 and was therefore possibly not identified before in studies dividing the VWFA based on functional differences. If we compare our results to another recent study (Boring et al., 2021), that also localized face areas (which may help localize the pFus word area), we find that in most of their subjects, a pFus word area (and the three other subareas) could be defined. Based on the split-half analysis, we demonstrated that these word selectivity areas in ventral OTC showed a higher response to words than to any of the other main categories. They also showed some response to other character categories, but a smaller response than to words.

We also located the mid-fusiform sulcus. The sulcus helped locate the face areas (Weiner et al., 2014), but it did not contribute much to locating the word areas. The split-half analysis showed that in these face areas, the response to faces was higher than to the other main categories. We also located several hand areas in the left hemisphere. One appeared lateral and posterior (corresponding to Bracci et al., 2010) and one more in the middle of the ventral OTC. The split-half analysis confirmed that these areas responded more to hands than to other categories. Nordt and colleagues (2021) suggested a competition for cortical territory between limb and word selectivity. In correspondence with that, we found that the hand areas also showed some selectivity to words. These hand areas acted as a reference point to locating the word and face areas: medial to the first hand area, appeared the first word and face area; in between the first and second hand area, there was a word area and medial to that one also the pFus face area (joined by the pFus word area); around the second hand area, there were the third word and face area. This result confirms the proposal of Yeatman and colleagues (2021), that the VWFA-1 and 2 would lie next to FFA-1 and FFA-2 and a limb-selective area separates the number 1 areas from the number 2 areas.

In the right hemisphere of OTC, the organization of selectivity was somewhat different than in the left hemisphere. We identified the first and second word area, but not the third (consistent with White et al., 2019), nor the pFus word area. The split-half analysis showed that this word selectivity was more generally a selectivity to characters (consistent with Dehaene et al., 2004; Vinckier et al., 2007). Face selectivity and the role of the mid-fusiform sulcus was similar to that of the left hemisphere. The first and second hand area could also be identified. The location of the word and hand areas was more complex in this hemisphere, and as a result hand areas could not serve as a reference point for locating

word and face areas. These findings are not surprising given the idea that the left hemisphere is the language dominant hemisphere.

In the representational space, we found that character categories grouped together, just like animate categories and inanimate categories respectively. The faces category, compared to hands and bodies, was not more similar to the words category, in terms of the neural pattern, as one might expect based on the neural recycling hypothesis (Dehaene & Cohen, 2007). This finding is in line with a recent study that shows the competition proposed by the neural recycling hypothesis may be between limbs and words (Nordt et al., 2021). Based on our finding, the competition with words may be with all human-related categories, like faces, bodies and hands. However, we also found that the neural pattern of faces was not more similar to words than the closest object category, cars. Based on this, it is possible that there is no competition between human-related categories and words at all in this space. Thus, the competition proposed by the neural recycling hypothesis may be limited to activity for categories and does not appear in activity patterns for categories. In the right hemisphere, the neural patterns for word and other categories were more similar, compared to the left hemisphere. This corresponds to our previous result where word (and other) selectivity was less specific in the right hemisphere.

The hand and tool areas

Previous research demonstrated the existence of hand and tool areas, both in the lateral and ventral left and right OTC (Bracci et al., 2010; Bracci et al., 2012; Bracci et al., 2016). To be able to locate all these hand areas, a lenient contrast between conditions was used: hands versus scrambled objects (Bracci et al., 2016) (in contrast to previous studies using hands versus chairs or animals, as in Bracci et al., 2010; Bracci et al., 2012). Evidence was also provided for an overlap between these areas, but it was noted that this overlap may not exist at the single neuron level and thus, may disappear when studied at a higher spatial resolution (Bracci et al., 2012). Previous research also used several spatial preprocessing steps that could influence the results (Weiner & Grill-Spector, 2013). To overcome these issues, we reinvestigated the hand and tool areas using 7T fMRI. More specifically, we studied the overlap between these areas in all parts of OTC, their response profiles and their representational spaces. We also defined face, body and object areas as a comparison.

A first important observation was that, using a strict contrast (hands or tools versus all other categories), we could locate all hand and tool areas. This suggested that we could, with our dataset, reinvestigate hand and tool selectivity, overcoming the drawbacks of previous studies. All four hand areas could distinguish between tools and other typical objects. They preferred tools over non-tool objects. Previous studies found this effect only in the left lateral hand area (Bracci et al., 2010; Bracci

et al., 2012; Bracci et al., 2016). The effect may have been smaller in these other hand areas and could therefore maybe only be picked up using 7T fMRI (which provides a higher signal-to-noise ratio), like in our study.

We found that the hand areas, except the one in left ventral OTC, responded more to bodies than to tools. This observation overall corresponded to previous research (Bracci et al., 2010; Bracci et al., 2012; Bracci et al., 2016). It is not surprising that hand areas would prefer bodies over tools as they are both animate categories and hands are a part of the body. We also, in contrast to previous research, included faces as a category. The lateral hand areas preferred tools over faces. The ventral hand areas responded to the two categories equally. This was surprising as faces is also an animate category like bodies, whereas tools are an inanimate category.

As mentioned above, we could also define the four tool areas in most participants. In right ventral OTC, this was more difficult: only eight participants had this area. This difficulty is consistent with a previous study (Bracci et al., 2012), although a later study did not have this issue, possibly due to using a lenient contrast (Bracci et al., 2016). The tool areas preferred hands above the other categories (faces, bodies, objects), although this effect was less clear in the right hemisphere. These findings are overall consistent with previous research (Bracci et al., 2012; Bracci et al., 2016).

In our study, we could also define face, body and object areas as a comparison. The face and body areas overall could not (in contrast to the hand areas), distinguish between tools and non-tool objects. Previous research did not define face areas. In one of them, a body area was defined and a similar test was also carried out, and their results are consistent with ours (Bracci et al., 2010). The object areas overall responded more to other typical objects than to tools. No previous research carried out a similar test so results cannot be compared.

The overlap between category-selective areas was also investigated. The overlap between hand and tool areas was larger than overlap between other category-selective areas and thus, we concluded that this overlap was specific to hands and tools. This corresponds to previous research (Bracci et al., 2012). The overlap was less specific in the right hemisphere, which cannot be compared to previous research as this was not tested before. This finding may however be a consequence of having more trouble defining the tool areas in this hemisphere. We also observed that the overlap was not larger in lateral versus ventral OTC. We did find that the overlap was larger in left lateral OTC than in right lateral OTC. We did not find the same effect for left versus right ventral OTC, but this could again be a consequence of the trouble defining tool areas in the right hemisphere, which was especially true for the right ventral OTC.

The representational spaces of the hand and tool areas were also investigated. In the hand areas, the neural representations of hands and tools were more similar than between hands and non-tool

objects. We could expect this based on the observation that hand areas prefer tools over non-tools. In the tool areas, the representations of hands and non-tool objects were equally similar to that of tools. This was not expected, as tool areas, based on previous research results (Bracci et al., 2010; Bracci et al., 2012; Bracci et al., 2016), either respond more to hands than to tools or respond equally high to them both. Previous research did not carry out multi-voxel pattern analysis with the same tests as we did, so we cannot compare our results.

The bigger picture

First, a clarification of the analysis choices in chapter 2 and 3 is warranted. In the second chapter, a split-half analysis was used to define regions of interest (e.g., word areas) and consequently analyze the functional profile of these areas. In the third chapter, all data was used to define regions of interest (e.g., hand areas) and consequently, using all the data, the functional profile of these areas was analyzed. To avoid the issue of circularity, we did not analyze the data of the category that was used to define the area (e.g., the functional response to hands was disregarded in the hand areas). The choice for a different analysis style in chapter 2 versus chapter 3 merits further clarification.

In chapter 2, the goal of the study was to define areas such as the word areas, to localize them (relative to other areas) and investigate the functional profile of these areas, in response to both the condition of interest for that area and other relevant categories of interest. One of the main reasons for this was the idea that word selectivity may in fact be a selectivity to symbols more in general. This idea requires to investigate the word versus other types of symbol selectivity within the identified word areas. Therefore, a split-half analysis was necessary.

In the third chapter of this thesis, our main goal was to confirm if indeed multiple lateral/ventral hand and tool areas exist in the occipitotemporal cortex and if so, if the hand and tool areas overlap significantly at this higher spatial resolution (without making certain analysis choices that could create this overlap). Therefore, we believed that using the full dataset was necessary, so that any lack of areas or overlap could not be due to using only half of the dataset in the analyses. As a result, when investigating the functional profile of these areas, one condition always had to be disregarded. However, because we did not have any clear hypotheses that would support the need for a split-half analysis (like we did in chapter 2), we believed that the benefits of not using such analysis outweighed the limitations it would introduce, in the context of our study. In an extra analysis, that was not included in the text of chapter 3, due to the already lengthy nature of the studies included in this thesis, we repeated the analyses of chapter 3 but then with a split-half analysis instead. The results of this analysis were along the same lines (with some tests no longer reaching statistical significance as is to be expected with only half of the data available for all tests) as the results reported in chapter 3, without using a split-half analysis.

Important to note is that for the multi-voxel pattern analysis, there is also a difference between the two chapters. In chapter 2, we used anatomically defined regions of interest, so that we could use the full dataset to understand the representational space, based on 20 different categories, of the whole of the occipitotemporal cortex. In chapter 3, we were however interested in the representational spaces of the various hand and tool areas specifically, and thus, we had to use a split-half analysis to define the regions of interest and consequently, use multi-voxel pattern analysis.

Second, applying to this thesis the concept of hierarchical processing, in the context of the visual system, is warranted. As explained in the general introduction of this thesis: two processing pathways, the ventral and the dorsal, start in posterior regions of the brain and then information flows further to anterior regions. This is a first more large-scale example of a processing gradient.

Such processing gradients can also be found within specific areas of the ventral pathway itself: the occipitotemporal cortex (OTC). Within OTC, there is also evidence for a processing gradient. Grill-Spector and Weiner (2014) described a processing gradient from lateral to medial in the ventral OTC. This processing gradient is followed by several organizational dimension of the OTC such as animacy: animate categories are processed more lateral and inanimate categories more ventral (Grill-Spector & Weiner, 2014). The mid-fusiform sulcus (MFS) acts as an anatomical boundary between what is lateral and ventral (Grill-Spector & Weiner, 2014). This distinction between lateral and ventral is further supported by underlying connectivity and microarchitecture (Grill-Spector & Weiner, 2014). In more detail, the authors propose that the lateral ventral OTC processes object-related info and has strong connectivity to the fusiform gyrus but less so to the parahippocampal gyrus, and the medial anterior part processes scene-related info and shows the opposite pattern of connectivity (Grill-Spector & Weiner, 2014).

Bao and colleagues (2020) suggest there may be hierarchical processing from posterior to anterior within OTC along areas supporting one specific category (e.g., faces): the areas may each provide different computations and interact with each other to end up with the full picture (of a face, e.g.). It has also been proposed that the face areas and the visual word form area (VWFA) may show such a processing gradient (Yeatman et al., 2021; Caffarra et al., 2021). This hypothesis is supported by evidence demonstrating different anatomical connectivity between VWFA-1 and VWFA-2 (Lerma-Usabiaga et al., 2018). Caffarra and colleagues (2021) elaborated on this proposal: the posterior VWFA processes perceptual info of written words, the anterior VWFA process linguistic aspects of the words. They also included temporal dynamics and anatomical connectivity evidence into their proposal (Caffarra et al., 2021). The word areas that we identified based on anatomical boundaries, correspond to previous functional subdivisions of the VWFA (e.g., Lerma-Usabiaga et al., 2018; White et al., 2019), and thus fit within this proposed model of posterior to anterior processing.

This processing gradient can also be applied to the hand and tool areas that we identified, as proposed by Bracci and colleagues (2016). The ventral areas in OTC provide category-specific object information, and the lateral areas in OTC provide a bridge between the visual information provided by ventral OTC and action information provided by parietal areas of the brain (Bracci et al., 2016). This proposal fits with functional connectivity results from a previous study that demonstrated a strong connectivity between lateral OTC and the intraparietal sulcus (Bracci et al., 2012). As far as I know,

there has been no direct evidence for the hypothesis on the connectivity of the ventral hand/tool areas with the lateral hand/tool areas of OTC or with parietal areas, which would be an interesting next step in this research field.

As shown by the results from the first chapter of this thesis, functional connectivity studies play an important part in understanding the functional organization of OTC. This becomes clear in this discussion on processing gradients, where connectivity is always a part of the theory. It is a limitation of this thesis that this analysis is not provided and this is discussed in more detail in the strengths and limitations section below. We can however expect that results of functional connectivity would confirm previous connectivity evidence because, as mentioned above, the word areas and the hand and tool areas that we identified corresponded to previously defined areas that were investigated in such connectivity analyses.

Strengths and limitations

Strengths and limitations specific to a certain chapter will be discussed first and then, more general strengths and limitations of this thesis will be considered.

The first chapter of this thesis compared for the first time, functional networks that could be built up from three methods of fMRI analyses. These networks were highly similar between the three methods, but each also contained some unique information that could not be predicted by the other methods' networks. Thus, each of these methods are important to investigate the functional organization of networks in the brain. With this study, researchers can be mindful of the differences between the methods and consequently, make an adequate choice between them based on their research goals.

This study did have several limitations. First, the data related to only one type of task: social touch scene perception. It is possible that the comparison between the methods is different in other tasks. We hypothesize that our conclusions will generalize to other task domains. Second, we only selected a limited number of regions of interest. We could have included instead a large number of network areas, but this was not possible due to the methods of the multivariate analysis. This analysis required meaningful task-related information to be present.

Regarding the second chapter of this thesis, it is important to note that the word forms used were letter strings, similar to the fake script and numbers categories, and were not pseudowords, nor did they have any semantic meaning. All words (except one) did contain at least one vowel and they could all be pronounced. We did find that the word areas of the left hemisphere were more responsive to words than to numbers and fake script, thus this is evidence that we identified at the minimum areas that differentiated between letters of the roman alphabet (compared to numbers and other letter types). Also, in the past, several different contrasts have been used to localize word selectivity, not just ones consisting of pseudowords or real words (see Caffarra et al., 2021 for an overview). As a result of the categories used, we could not make functional distinctions between word areas, we could only differentiate between them based on anatomical boundaries. However, the found areas do correspond to previously reported functionally divided subareas (Lerma-Usabiaga et al., 2018; White et al., 2019). Another limitation of this study was that, although across subjects we could identify a typical organization of word selectivity, this organization was not as clear in each of the subjects. This is to be expected though, based on individual variability.

To our knowledge, this study is the first to map, with a high spatial resolution at the level of the individual brain, word selectivity and the anatomical boundaries between the subareas of the VWFA, in combination with mapping other types of category selectivity. This answered the need pressed by recent prominent researchers in the field (Caffarra et al., 2021; Yeatman et al., 2021). These results

will enable future research to define the VWFA and its subareas consistently. On top of that, we demonstrated the advantage of including selectivity to hands in a study to reliably localize word (and face) areas. This study also showed that the popular idea of competition between words and faces, from the neural recycling hypothesis, may extend to other categories such as hands (in line with Nordt et al., 2021). We also showed that the competition idea does not extend to the representational space of OTC.

In the third chapter of this thesis, we were able to employ a high spatial resolution and omit several spatial preprocessing steps that may have an influence on the results regarding the hand and tool areas. This influence could be as big as not finding an overlap between these areas anymore and was thus an important next step to take in this field. We demonstrated that this overlap remained. We also showed that all hand areas distinguished between tools and other objects, and not just the left lateral hand area. In our study, we also investigated other human categories like bodies and faces. By doing this, we showed for the first time that the hand areas had a higher response to tools than to faces. On top of that, the results showed that all tool areas and not just the left lateral tool area, could distinguish between hands, faces and non-tool objects. This study was also the first to investigate the overlap between hand and tool areas beyond the left lateral OTC (as carried out by Bracci et al., 2012). The overlap was specific to hands and tools in the left hemisphere, and less so in the right hemisphere (i.e., there was also overlap between other categories). We also for the first time, constructed the representational space of all the hand and tool areas and compared the distance between neural patterns for multiple pairs of categories. In our study, we also ensured that body stimuli did not contain hands. They do not contain faces as that is a separate category, thus they should also not contain hands. Although this may not have influenced results in previous studies, we believe this experimental detail good practice, to use also in future studies, because the effect this detail has on results is unclear. Notably, this study also had its limitations. We could not define certain category-selective areas in many of the participants. This was especially true in the univariate analyses. This influenced all tests concerning these areas (either relating to the response profiles or the overlap between areas).

I will now focus on more general strengths and limitations of this PhD thesis.

One limitation of the studies in the second and third chapter is that no functional connectivity analysis was carried out. This is a limitation given that without this analysis, it is not possible, as shown by the results of the study in the first chapter, to make conclusions about how the category-selective areas of interest form networks both within and outside of OTC. We opted to focus first on univariate and multivariate analyses because these analyses could much better answer our main research questions. Due to time constraints, it was not possible to carry out this extra analysis of functional connectivity as well.

Another limitation concerns the definition of contrasts in the univariate analyses, which are crucial for identifying selectivity to specific categories and merit careful consideration. In our studies in chapter 2 and 3, we aimed to balance between the need for specificity and practicality when defining contrasts. Consequently, we refrained from employing contrasts such as one category (e.g., faces) versus scrambled images or versus the cubies and smoothies categories. These contrasts would risk capturing selectivity to shapes (with a semantic meaning) in general, and are thus not specific enough. Alternatively, we could have opted for a highly specific contrast, such as faces versus all animate categories, ensuring a focus on pure face selectivity rather than broader animacy selectivity. However, this makes it challenging to detect any selectivity. Thus, we chose a more practical strategy that would also allow to capture a moderately specific selectivity for each category of interest: contrasting each category against all other categories individually. It is important to acknowledge a slight bias inherent to this type of contrast. Contrasting against all other categories introduces a larger number of inanimate than animate categories into the comparison. Consequently, when contrasting an animate category like faces versus all other categories, detecting selectivity to faces might be more challenging. Conversely, when contrasting an inanimate category like chairs versus all other categories, the bias may make it easier to identify selectivity to chairs. We believe that this potential bias is minimal and somewhat mitigated by the substantial size of the 'all other categories' group, consisting of 19 categories. It is noteworthy that among our 20 categories, five are clearly animate, six are clearly inanimate, and certain categories, such as words, smoothies, and flowers, are not so straightforward to define on the animacy dimension (as they are symbols, unknown, or living shapes). Therefore, it is challenging to definitively conclude that an animacy bias exists in our contrasts.

A strength of the two studies of chapter 2 and 3 is that they use 7T fMRI, which provided several benefits to the topics of these studies. First, it provides a high spatial resolution and a high-quality signal. This is confirmed by, for example, in the results of chapter 2, we observed that activity was restricted to the grey matter. This is also confirmed by, the third chapter, where we found that we could locate all hand areas with a very strict contrast between conditions (in contrast to previous research). This also allowed us to investigate if the overlap between hand and tool areas remained at a higher resolution. Second, 7T fMRI allowed us to conduct analyses at the level of the individual brain, without spatial normalization and smoothing of the data. This is a benefit for several reasons. The data stays true to the gyral and sulcal anatomy of the brain, which then allows a more accurate localization of activity (Weiner & Grill-Spector, 2013). Spatial smoothing can also lead to averaging together areas that actually lie separate on the brain surface (Weiner & Grill-Spector, 2013). Group averaging results across participants can lead to a lack in spatial precision and by this, boundaries between areas can become blurred, especially if areas vary in size and/or location across participants (Wandell et al.,

2012; Glezer & Riesenhuber, 2013). These factors were crucial for both chapter 2 and 3. For the second chapter for example, it was important to localize activity accurately so we could find a consistent organization across subjects. For the third chapter for example, the overlap between areas could have been a result of smoothing or group averaging. Third, 7T fMRI, through its high-quality signal, enabled us to use a large number of categories in the experiments of the second and third chapter, which will be discussed in the next paragraph in more detail.

Previous research has shown that several dimensions play a role in the functional organization of OTC, such as animacy (Grill-Spector & Weiner, 2014) and real-world size (Konkle & Oliva, 2012). Therefore, it is necessary to include a large number of categories when investigating OTC's organization. A study in 2006 by Downing and colleagues attempted this by using 20 categories. Due to the use of 1.5T fMRI however, they suffered from low spatial resolution and signal of a lower quality compared to 3T and certainly compared to 7T. They also did not employ multi-voxel pattern analysis (MVPA). Other studies often focus on a small number of categories and then may or may not use MVPA (e.g., Weiner & Grill-Spector, 2010; Bracci et al., 2016), or studies use a large number of categories and focus MVPA, disregarding detailed univariate analyses (e.g., Kriegeskorte et al., 2008). In the study of chapter 2 and 3, we included many categories that vary on several different dimensions: animacy, natural/artificial, real-world size, characters, etcetera. This is possible due to the high-quality signal that 7T fMRI provides. We focused on both univariate analysis and MVPA. This is an important strength of the 7T fMRI study of the second and third chapter.

During the creation of the stimuli used in the second and third chapter, we made the choice between either natural images or images controlled for differences in low-level features between categories (e.g., the color is made similar between the images of the different categories). We chose for the second option. Therefore, these studies are somewhat limited in generalizability because the stimuli are manipulated in several ways and thus less similar to objects we encounter in real life. The advantage of this choice however is for the objectives of the studies. In these studies, we wanted to make implications about category-selectivity specifically: if an area was selective to trees, we wanted to be able to conclude that this selectivity was related to trees specifically and not for example to the green color it typically has, that the other categories did not have.

Future research

Regarding the first chapter, future studies could compare networks from different methods and investigate if the size and the number of regions of interest included in the networks, influence the conclusions regarding the similarity of results between the different methods of analyses. Future research could also investigate if our conclusions hold in other task contexts. Another interesting investigation would compare more than just these three networks. For example, multivariate functional connectivity is another relevant method to compare to (Coutanche & Thompson-Schill, 2014; Anzellotti & Coutanche, 2018).

Related to the VWFA, the functional connectivity of the subareas of the VWFA as we defined them, with other brain areas, could be investigated in the future. There has been ample research on anatomical connectivity (see Caffarra et al., 2021; Yeatman et al., 2021, for an overview), but functional connectivity research could add new insights into the roles of the VWFA and its subareas specifically (Caffarra et al., 2021). Another path for future research is investigate the functional differences between the subareas of the VWFA further. Based on the correspondence between the subareas as we defined them anatomically, and the functional subdivisions made by Lerma-Usabiaga (2018) and White and colleagues (2019), we can already form some idea of the functional differences between them. More research is needed on this topic and should also include the pFus word area. For example, the posterior-to-anterior gradient (explained in the section on the bigger picture) can be further investigated. Future research could also carry out a similar study as ours that also includes pseudowords and real words and by this, can investigate functional differences between subareas. This could also be interesting for investigating word selectivity in the representational space of OTC. Future research could also investigate other categories than just words, faces and hands, both in the context of research on the competition from the neural recycling hypothesis and in the context of the location and function of areas of category selectivity.

Regarding the hand and tool areas, future research could investigate all the hand and tool areas separately in multiple task conditions to understand their specific functions. Research with an even higher spatial resolution could also be important because it is still possible that at such a resolution, the apparent overlap between hand and tool areas disappears because there may in reality not be overlap at the level of the single neuron. Finally, research could investigate the functional connectivity of these hand and tool areas with each other and with other brain areas, which may differ depending on the specific hand/tool area. Research on the functional connectivity has been limited, as far as I know, on connectivity between lateral areas and parietal areas (Bracci et al., 2012). Extending this research to other types of connectivity would aid the understanding of the functions of these areas both within OTC and in the rest of the brain.

More generally, supported by the results of the hand areas, future research could employ the benefits of 7T fMRI to research if in OTC, there are category-selective areas that follow the so-called homunculus. This is a somewhat distorted human figure that signifies how different parts of the body are being processed, both in the motor system and in the sensory system of the brain. Body parts that require fine processing are represented more in those respective brain systems than body parts requiring less precise processing. For example, more neurons of those systems are devoted to the hands and the face than to the legs or the trunk. Bratch and colleagues (2018) provided some first evidence in support of this idea.

Another idea for future studies would be to provide laminar and/columnar information on the organization of OTC. For this, one would need ultra-high spatial resolution. This would allow for example to compare columnar results on the organization of OTC in primates with humans (Tanaka, 2003). It can also allow to understand the role of the superficial, middle and deep layers of the cortex, which is gradually being uncovered in other fields related to the human brain (see Zaretskaya, 2021, for a review). One recent study has attempted a laminar analysis of the ventral OTC (Margalit et al., 2020). They found the superficial layers of the cortex in this location to respond to a wide range of stimuli and visual features and to represent domain-general info, but the deeper into the layers, the more specific the response becomes (to specific features such as faces) and the more domain-specific the representations are (Margalit et al., 2020).

References

- Anzellotti, S., & Coutanche, M. N. (2018). Beyond Functional Connectivity: Investigating networks of multivariate representations. *Trends in Cognitive Sciences*, 22(3), 258–269.
<https://doi.org/10.1016/j.tics.2017.12.002>
- Bao, P., She, L., McGill, M., & Tsao, D. Y. (2020). A map of object space in primate inferotemporal cortex. *Nature*, 583(7814), 103–108. <https://doi.org/10.1038/s41586-020-2350-5>
- Boring, M. J., Silson, E. H., Ward, M. J., Richardson, R. M., Fiez, J. A., Baker, C. I., & Ghuman, A. S. (2021). Multiple adjoining word- and face-selective regions in ventral temporal cortex exhibit distinct dynamics. *The Journal of Neuroscience*, 41(29), 6314–6327.
<https://doi.org/10.1523/JNEUROSCI.3234-20.2021>
- Bracci, S., Cavina-Pratesi, C., Connolly, J. D., & Ietswaart, M. (2016). Representational content of occipitotemporal and parietal tool areas. *Neuropsychologia*, 84, 81–88.
<https://doi.org/10.1016/j.neuropsychologia.2015.09.001>
- 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. <https://doi.org/10.1152/jn.00619.2011>
- 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. <https://doi.org/10.1152/jn.00215.2010>
- Bratch, A., Vizioli, L., Yacoub, E., Engel, S. A., Burton, P. C., & Kersten, D. J. (2018). Selectivity for individual body parts within body selective visual cortex. Poster presented at OHBM 2019, Rome, Italy.
- Caffarra, S., Karipidis, I. I., Yablonski, M., & Yeatman, J. D. (2021). Anatomy and physiology of word-selective visual cortex: from visual features to lexical processing. *Brain Structure & Function*, 226(9), 3051–3065. <https://doi.org/10.1007/s00429-021-02384-8>
- Coutanche, M. N., & Thompson-Schill, S. L. (2014). Using informational connectivity to measure the synchronous emergence of fMRI multi-voxel information across time. *Journal of Visualized Experiments*, 89. <https://doi.org/10.3791/51226>
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J.-B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence: Behavioral and neuroimaging evidence. *Psychological Science*, 15(5), 307–313.
<https://doi.org/10.1111/j.0956-7976.2004.00674.x>
- Dehaene, Stanislas, & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56(2), 384–398.
<https://doi.org/10.1016/j.neuron.2007.10.004>

- Downing, P. E., Chan, A. W.-Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cerebral Cortex*, 16(10), 1453–1461. <https://doi.org/10.1093/cercor/bhj086>
- Glezer, L. S., & Riesenhuber, M. (2013). Individual variability in location impacts orthographic selectivity in the “visual word form area.” *The Journal of Neuroscience*, 33(27), 11221–11226. <https://doi.org/10.1523/JNEUROSCI.5002-12.2013>
- Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, 15(8), 536–548. <https://doi.org/10.1038/nrn3747>
- James, K. H., James, T. W., Jobard, G., Wong, A. C. N., & Gauthier, I. (2005). Letter processing in the visual system: different activation patterns for single letters and strings. *Cognitive, Affective & Behavioral Neuroscience*, 5(4), 452–466. <https://doi.org/10.3758/cabn.5.4.452>
- Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron*, 74(6), 1114–1124. <https://doi.org/10.1016/j.neuron.2012.04.036>
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2, 4. <https://doi.org/10.3389/neuro.06.004.2008>
- Lee Masson, H., Van De Plas, S., Daniels, N., & Op de Beeck, H. (2018). The multidimensional representational space of observed socio-affective touch experiences. *NeuroImage*, 175, 297–314. <https://doi.org/10.1016/j.neuroimage.2018.04.007>
- Lerma-Usabiaga, G., Carreiras, M., & Paz-Alonso, P. M. (2018). Converging evidence for functional and structural segregation within the left ventral occipitotemporal cortex in reading. *Proceedings of the National Academy of Sciences of the United States of America*, 115(42), E9981–E9990. <https://doi.org/10.1073/pnas.1803003115>
- Margalit, E., Jamison, K. W., Weiner, K. S., Vizioli, L., Zhang, R.-Y., Kay, K. N., & Grill-Spector, K. (2020). Ultra-high-resolution fMRI of human ventral temporal cortex reveals differential representation of categories and domains. *The Journal of Neuroscience*, 40(15), 3008–3024. <https://doi.org/10.1523/JNEUROSCI.2106-19.2020>
- 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. <https://doi.org/10.1093/scan/nsn044>
- Nordt, M., Gomez, J., Natu, V. S., Rezai, A. A., Finzi, D., Kular, H., & Grill-Spector, K. (2021). Cortical recycling in high-level visual cortex during childhood development. *Nature Human Behaviour*, 5(12), 1686–1697. <https://doi.org/10.1038/s41562-021-01141-5>

- Pillet, I., Op de Beeck, H., & Lee Masson, H. (2019). A comparison of functional networks derived from representational similarity, functional connectivity, and univariate analyses. *Frontiers in Neuroscience*, 13, 1348. <https://doi.org/10.3389/fnins.2019.01348>
- Raizada, R. D. S., Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2010). Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: prediction of individual differences. *Cerebral Cortex*, 20(1), 1–12. <https://doi.org/10.1093/cercor/bhp076>
- Strother, L., Coros, A. M., & Vilis, T. (2016). Visual cortical representation of whole words and hemifield-split word parts. *Journal of Cognitive Neuroscience*, 28(2), 252–260. https://doi.org/10.1162/jocn_a_00900
- Tanaka, K. (2003). Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cerebral Cortex*, 13(1), 90–99. <https://doi.org/10.1093/cercor/13.1.90>
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron*, 55(1), 143–156. <https://doi.org/10.1016/j.neuron.2007.05.031>
- Wandell, B. A., Rauschecker, A. M., & Yeatman, J. D. (2012). Learning to see words. *Annual Review of Psychology*, 63(1), 31–53. <https://doi.org/10.1146/annurev-psych-120710-100434>
- Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., Amunts, K., & Grill-Spector, K. (2014). The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage*, 84, 453–465. <https://doi.org/10.1016/j.neuroimage.2013.08.068>
- Weiner, K. S., & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage*, 52(4), 1559–1573. <https://doi.org/10.1016/j.neuroimage.2010.04.262>
- 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. <https://doi.org/10.1007/s00426-011-0392-x>
- White, A. L., Palmer, J., Boynton, G. M., & Yeatman, J. D. (2019). Parallel spatial channels for word recognition converge at a bottleneck in anterior word-selective cortex. *Journal of Vision*, 19(10), 173a. <https://doi.org/10.1167/19.10.173a>
- Wong, A. C.-N., Jobard, G., James, K. H., James, T. W., & Gauthier, I. (2009). Expertise with characters in alphabetic and nonalphabetic writing systems engage overlapping occipito-temporal areas. *Cognitive Neuropsychology*, 26(1), 111–127. <https://doi.org/10.1080/02643290802340972>

- Xue, S., Weng, X., He, S., & Li, D. (2013). Similarity representation of pattern-information fMRI. *Chinese Science Bulletin*, 58(11), 1236–1242. <https://doi.org/10.1007/s11434-013-5743-0>
- Yeatman, J. D., & White, A. L. (2021). Reading: The confluence of vision and language. *Annual Review of Vision Science*, 7(1), 487–517. <https://doi.org/10.1146/annurev-vision-093019-113509>
- Zaretskaya, N. (2021). Zooming-in on higher-level vision: High-resolution fMRI for understanding visual perception and awareness. *Progress in Neurobiology*, 207(101998), 101998. <https://doi.org/10.1016/j.pneurobio.2021.101998>
- Zeharia, N., Hertz, U., Flash, T., & Amedi, A. (2015). New whole-body sensory-motor gradients revealed using phase-locked analysis and verified using multivoxel pattern analysis and functional connectivity. *The Journal of Neuroscience*, 35(7), 2845–2859. <https://doi.org/10.1523/JNEUROSCI.4246-14.2015>

General Conclusion

As a first objective of this PhD, we aimed to compare three popular analyses of fMRI data so that future research could make an informed choice between them. We provided such a study in the first chapter and showed that functional networks derived from these three methods of analysis are highly similar, but also contain specific information that cannot be predicted by another method. Therefore, future research should consider its research questions and goals and ensure that the chosen methods of analysis fit, because this choice determines which implications can be derived from a dataset.

A second objective of this PhD was to identify and investigate subareas of the visual word form area, and relate them to other category-selective areas, using a high spatial resolution at the level of the individual brain. We identified three subareas of the visual word form area, which correspond to previous evidence, and one extra subarea that was joined by a face area, which we termed the pFus word area. We also uncovered that hand areas could provide a reference point from which word (and face) areas could be located. By investigating the representational space, we could also conclude that a competition between words and faces does not exist in this space.

A third objective of this PhD was to study the hand and tool areas and their overlap. We could identify all four hand and tool areas. We demonstrated that all hand areas differentiated between tools and non-tool objects, and that all tool areas differentiated between hands and other categories such as faces. We provided evidence that the overlap between hand and tool areas was specific (the overlap between these areas was significantly higher than overlap between other category-selective areas). The representational space of the hand and tools areas was also studied. From this, we could conclude that in the hand areas, neural representations for hands and tools were more similar than for hands and non-tool objects. In the tool areas, representations for hands, tools and non-tool objects were equally similar to one another.

In general, this PhD aimed to provide a more detailed understanding of the neural basis for visual categorization. We achieved this by first, providing a methodological study about three popular analyses for fMRI data of both the occipitotemporal cortex and the brain more in general. Second, we provided a 7T fMRI study on the VWFA that has uncovered more details on findings of previous research (such as the anatomical location of subdivisions in relation to other category-selective areas) and provided new findings (such as the finding of a pFus word area). Third, by investigating the hand and tool areas using 7T fMRI, we showed that they could all be identified with a stricter contrast compared to previous research and that not only the lateral but also the ventral areas showed overlap between them. Thus, we again provided more details on findings from previous research, but we also added new findings to this field of topic, by for example adding the faces category to the investigation.