

Imaging imagery

Citation for published version (APA):

Emmerling, T. C. (2016). *Imaging imagery: an investigation of visual cognition using high-resolution fMRI*. [Doctoral Thesis, Maastricht University]. CPI Wöhrmann Print Services B.V.
<https://doi.org/10.26481/dis.20160622te>

Document status and date:

Published: 01/01/2016

DOI:

[10.26481/dis.20160622te](https://doi.org/10.26481/dis.20160622te)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

IMAGING IMAGERY
An investigation of visual cognition using
high-resolution fMRI

THOMAS CHRISTOPH EMMERLING

Thomas Christoph Emmerling: *Imaging imagery*, An investigation of visual cognition using high-resolution fMRI, © June 2016

No part of this publication may be reproduced, stored in an automated data system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording or otherwise, without prior permission of the author.

Printed by CPI-Koninklijke Wöhrmann – Zutphen

ISBN: 978-94-6328-049-5

IMAGING IMAGERY

An investigation of visual cognition using high-resolution fMRI

THOMAS CHRISTOPH EMMERLING

DISSERTATION

to obtain the degree of Doctor at Maastricht University,
on the authority of Rector Magnificus, Prof. Dr. L.L.G. Soete,
in accordance with the decision of the Board of Deans,
to be defended in public on

Wednesday 22nd of June 2016 at 10.00 hours

by

Thomas Christoph Emmerling

SUPERVISOR:

Prof. Dr. Rainer Goebel

CO-SUPERVISOR:

Dr. Martin Frost

ASSESSMENT COMMITTEE:

Prof. Dr. Elia Formisano (Chair)

Prof. Dr. Lars Muckli, University of Glasgow, United Kingdom

Prof. Dr. Alexander Sack

Dr. Marieke Mur, University of Cambridge, United Kingdom

The work in this thesis was supported by grants from the European Research Council (ERC; Grant Agreement Number: 269853).

CONTENTS

1	GENERAL INTRODUCTION	1
2	OVERLAPPING FEATURE REPRESENTATIONS IN HUMAN VISUAL CORTEX	21
3	DECODING THE DIRECTION OF IMAGINED MOTION	63
4	RECONSTRUCTING AND DECODING IMAGINED LETTERS FROM THE VISUAL SYSTEM	115
5	GENERAL DISCUSSION	151
	Summary	169
	Valorisation	173
	Acknowledgements	179
	Curriculum Vitæ	185
	Publications	187

Chapter 1

General Introduction

Vision is one of the most important senses in humans. We use it every day, not only to navigate our environment but also as a vehicle for more complex cognitions: Many things we do on a daily basis – reading, writing, experiencing a piece of art, or interpreting the emotion of another individual – rely on input perceived with our eyes and on mental models of objects that we saw and remember. Vision holds crucial information that we make use of everyday, whether consciously or unconsciously.

1.1 EARLY THEORIES

Due to the obvious importance of vision for humans, an early interest in how we perceive the world with our eyes (A. M. Smith, 2014) arose. Plato tried to explain the phenomenon of visual perception around 360 BC in his dialogue *Timaeus* postulating that the eyes emitted light rays that enabled humans to see (*emission theory*). Other theories – as championed by Epicurus around 300 BC – opposed this view in that they saw the eyes as a detector for particles that are reflected from the objects that we perceive (*intromission theory*). The latter perspective comes closer to our modern understanding and was further supported by more modern philosophers and scientists like Isaac Newton and John Locke in the 18th century. It is worth mentioning, however, that recent studies found about 50% of adults to still (today) carry beliefs close to emission theory (Gregg, Winer, Cottrell, Hedman, & Fournier, 2001; Winer, Cottrell, Gregg, Fournier, & Bica, 2002).

Modern research on human vision was pioneered by scientists such as Thomas Young (1773–1829) and Hermann von Helmholtz (1821–1894) who hypothesised about the different types of photoreceptors in the human eye that allow us to perceive colours

(Young, 1802). Early theories on human vision focussed on the eye, as it is the most obvious location in which to search for the implementation of human vision. However, these early attempts could neither locate nor explain the computational operations that are necessary to segment an image into objects and, in the end, synthesise our final visual percept. Human visual perception remained elusive and a real understanding only set in with modern neuroscience.

1.2 NEUROSCIENCE OF VISION

Today, the visual system is one of the best-studied parts of human neocortex. Alongside the huge interest for how human visual perception is implemented in the brain, there are several reasons why modern neuroscience has invested so much energy investigating the visual system. First, the possibility of an easy external visual stimulation by means of, for instance, a computer monitor allows well-controlled experiments that can be carried out over extensive periods of time. This is important for the successful application of typical neuroimaging methods like electroencephalography (EEG) or functional magnetic resonance imaging (fMRI) as their signal-to-noise ratio (SNR) is low and, hence, experimental designs rely mostly on the averaging of data. An averaging approach calls for well-controlled experimental conditions including many data points. Second, the cortical surface area covered by the visual system, is large (55 % of the macaque neocortex' surface; Felleman and Van Essen, 1991) in comparison to other primary sensory areas like somatosensory (11 %) or auditory cortex (3 %). This allows investigations using neuroimaging techniques that are of sub-optimal spatial resolution, like EEG or standard-resolution fMRI.

Third, the functional organisational principles of the visual system can be studied in other species, often leading to subsequent findings in homologue areas of human cortex (e. g. Albright, 1984; Zimmermann et al., 2011). This is particularly important as animal studies allow more invasive neuroimaging techniques than studies in humans. Through comparison between the organisational principles of the visual system in different species, computational hierarchies can be understood on a more generalised level. Finally, lesions to cortical areas encompassing the visual system lead to very specific behavioural deficits, like hemianopsia (the loss of sight in one half of the visual field), achromatopsia (a deficit in colour perception), akinetopsia (motion blindness), or prosopagnosia (impaired ability to recognise faces). Due to the cortical size of the visual system and specifically testable visual deficits, lesion studies led to early important insights into human vision. In fact, lesion studies are the historical beginning of modern neuroscience (Rorden & Karnath, 2004).

1.3 COARSE ORGANISATION OF THE VISUAL SYSTEM

The visual system is mainly located in the occipital lobe extending into the parietal and temporal lobes of the human brain (see also [Figure 1.1](#)). When light rays are reflected from an object in front of us and enter our eyes, they hit receptors on the retina that detect the incoming photons and produce an action potential. The photoreceptors on the retina are connected through the optical nerve to the lateral geniculate nucleus (LGN) of the thalamus, which in turn projects to primary visual area V_1 located around the calcarine sulcus in the left and right hemisphere, also called striate cortex. From V_1 , connections project to extrastriate areas such as

second visual area V_2 and third visual area V_3 (both located around V_1) and then further on to areas V_4 and V_5 (the latter also called medial temporal area or MT). In total, the visual system encompasses more than 30 distinct areas (Van Essen, Anderson, & Felleman, 1992). Although many of these areas are organised in a hierarchy, with information flowing from lower (e.g. V_2) to higher (e.g. V_4) areas, additional feedback loops and interconnections culminate in more than 300 pathways connecting visual areas (Van Essen et al., 1992). This renders the visual cortex a complicated structure but also provides the basis for its remarkable abilities and dynamics.

The definition of subareas within the visual system is still an ongoing process, especially with respect to higher level areas, and historically relies on the usage of several neuroscientific methods. Cytoarchitectonic differences allow one to distinguish Brodmann areas (BAs) (Zilles & Amunts, 2010) that form the microanatomical definition of the functional areas V_1 (BA17), V_2 (BA18), and V_3 (BA19). In other areas of the visual system, anatomical variations correspond to and, thus, enable the identification of underlying functional units (Weiner & Zilles, 2015).

Results from functional neuroimaging methods are crucial for visual neuroscience. Starting from electrophysiological studies in the cat and monkey cortex, neuroscientists began to understand the functional architecture of visual cortex on a finer level. When in 1959 Hubel and Wiesel discovered small-scale functional tunings for spatial orientation in the cat's cortex (Hubel & Wiesel, 1959, 1962), they introduced the concept of systematic cortical organisation with neurons grouped into (hyper-)columns (Hubel & Wiesel, 1968; Mountcastle, 1997, 1957). This seminal discovery was one of the first steps towards a deeper understanding of not only organisational principles of visual cortex, but also computational

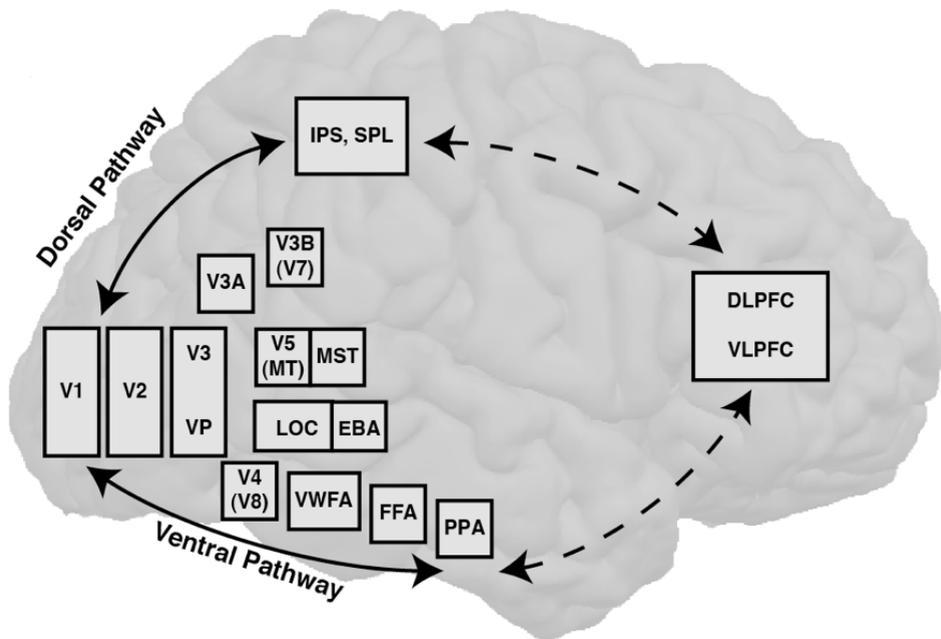


Figure 1.1 Prominent cortical areas in the human visual system. Connections (both feedforward and feedback) exist between many of these areas. Main dorsal and ventral pathways are shown (solid lines) including assumed pathway extensions into prefrontal cortex (dashed lines). MT: middle temporal area; MST: medial superior-temporal areas; LOC: lateral occipital cortex; FFA: fusiform face area; PPA: parahippocampal place area; IPS: intraparietal sulcus; SPL: superior parietal lobule; DLPFC: dorso-lateral prefrontal cortex; VLPFC: ventro-lateral prefrontal cortex. (based on Goebel, Muckli, and Kim, 2011)

steps undertaken by the brain to process incoming information from the retina.

Ever since, visual neuroscience set out to map the functional properties of cortical areas in the visual system and discovered specialisations for, e. g., motion (Tootell et al., 1995), colour (McKeefry & Zeki, 1997), orientation (Tootell et al., 1998), or binocular disparity (Backus, Fleet, Parker, & Heeger, 2001). More complex features like a preference for faces (Kanwisher, McDermott, & Chun, 1997), places (Epstein & Kanwisher, 1998), or objects (Kourtzi & Kanwisher, 2001) could be associated to hierarchically higher cortical areas in the visual system. Often, these specialisations were first found in animal models and later replicated in homologue areas of human cortex using non-invasive methods. A very prominent example is retinotopy, a technique pioneered by Sereno et al. (1995), that allows mapping the early visual areas of human cortex using fMRI and a finely tuned visual stimulation. Given the level of detail for results from animal studies, human neuroscience often lagged behind in mapping specialisations for visual features at equal accuracy and spatial resolution.

Over the last decades, neuroscience has revealed numerous details about the functional architecture of and the computations taking place along the hierarchically organised processing streams within the human visual system. Still, the fine-grained organisation of functional areas remained elusive. This deficit is related to the hitherto existing limitations regarding the spatial resolution of non-invasive neuroimaging techniques. In the next sections, the recent developments in fMRI methodology that can help to overcome this limitation will be discussed. Furthermore, animal models are useful for understanding relatively simple forms of visual cognition. When it comes to complex tasks in which humans outperform animals or even tasks that only our species can perform,

we have to investigate the visual cortex of humans in order to understand the computations underlying these abilities. Oftentimes, such tasks involve our implicit or explicit knowledge about the world which influences our visual system to accomplish the tasks. This calls for investigations of top-down influences in the visual system during which activity in higher order brain areas in temporal, parietal, and frontal cortex change the neural activation patterns in cortical areas along the visual hierarchy. One specific form of top-down influence is visual mental imagery and will be discussed in the following sections.

1.4 ADVANCES IN FMRI

Shortly after its inception in 1990 (Kwong et al., 1992; Ogawa & Lee, 1990) fMRI became the workhorse of cognitive neuroscience in general and human visual neuroscience in particular (Bandettini, 2012). It provided an unprecedented level of spatial resolution and depth of imaging while being non-invasive (compared to other imaging techniques like positron emission tomography). While MRI made it possible to image different tissue types within the brain in vivo, by measuring their respective influence on the relaxation of proton spins, it was the discovery of the blood-oxygen level dependent (BOLD) response that made the brain's functional processes visible. Here, the magnetic susceptibility differences between oxygenated and deoxygenated blood that influence the MRI signal are employed to estimate the local influx of oxygen-rich blood in each voxel (three-dimensional pixel) for multiple time points. This oxygen level is coupled to neuronal activity and, thus, reflects functional activation of the brain, although the precise mechanisms of this neurovascular coupling are still un-

der debate (Hirsch, Reichold, Schneider, Székely, & Weber, 2012; Uludağ, Müller-Bierl, & Uğurbil, 2009).

In recent years, methodological advancements in the field of fMRI have enabled us to reveal the functional organisations of cortical areas at a finer level. Going to magnetic field strengths of 7 Tesla (T) and beyond (Harel, 2012; Uğurbil et al., 2003; Yacoub et al., 2001) allows imaging at functional resolutions of 1 mm isotropic voxel sizes and below. At this mesoscopic resolution, it becomes feasible to unravel columnar structures in cortical areas (Cheng, Waggoner, & Tanaka, 2001; Goncalves et al., 2015; Yacoub, Harel, & Uğurbil, 2008; Yacoub, Shmuel, Logothetis, & Uğurbil, 2007; Zimmermann et al., 2011). Moreover, higher magnetic field strengths yield increased SNRs (Vaughan et al., 2001) and contrast-to-noise ratios (CNRs) over conventional 3 T fMRI, especially when combined with advanced fMRI sequences (Feinberg, Harel, Ramanna, Uğurbil, & Yacoub, 2008; Feinberg & Yacoub, 2012; Shmuel, Yacoub, Chaimow, Logothetis, & Uğurbil, 2007). Modern spin-echo sequences seem to increase sensitivity to BOLD responses originating in local microvasculature even further (De Martino et al., 2013; Uğurbil et al., 2003; Uludağ et al., 2009; Yacoub et al., 2007), while the influence of large draining veins on the BOLD response is diminished. Together, these developments pushed the limits of non-invasive neuroimaging techniques that can be employed to analyse and understand the functional organisation in the human (visual) cortex.

1.5 TOP-DOWN INFLUENCES AND IMAGERY IN THE VISUAL SYSTEM

Most previous studies on the visual system have focussed on perception and have not explicitly investigated other types of activations and normal-day usage of visual cortex. Beyond mere bottom-up perception, top-down influences, accounting for a large part of how we perceive the world, are often underestimated (Gilbert & Sigman, 2007). Recent studies found brain activity in early visual areas in the absence of actual visual stimulation, due to contextual feedback on a stimulus from higher visual areas (Muckli et al., 2015; F. W. Smith & Muckli, 2010) or even auditory areas (Vetter, Smith, & Muckli, 2014). This sheds light on the importance of top-down influences occurring even at the lowest levels of the visual hierarchy.

A special case of top-down influences on the visual system is imagery. Visual mental imagery contributes to human creativity and is essential for our ability to solve problems (Ferguson, 1977; Kozhevnikov, Kozhevnikov, Yu, & Blazhenkova, 2013). Imagination is at the core of many processes in the human mind (Miller, 1996). Unravelling the coding of mental imagery within the brain is, thus, of utmost importance for the advance of cognitive neuroscience. There is a long-standing debate about the neurocognitive implementation of mental imagery (Kosslyn, Thompson, & Ganis, 2006; Tye, 2000): It is debated, whether there is an internal picture that is generated and can be “seen with the mind’s eye” (e.g. Kosslyn, 1980) or whether such images would be epiphenomenal and people would mainly rely on a mere description and tacit knowledge of an imagined stimulus when solving imagery tasks (e.g. Pylyshyn, 1973). Answering this question has proven to be difficult because of the inherently private nature of and

inter-individual differences in imagery (Kozhevnikov, Hegarty, & Mayer, 2002; Kozhevnikov, Kosslyn, & Shephard, 2005; Poltrock & Brown, 1984) and the variability of brain areas that individuals employ during imagery (Kosslyn, Ganis, & Thompson, 2001; Motes, Malach, & Kozhevnikov, 2008; Wraga, Thompson, Alpert, & Kosslyn, 2003). This led to a comparative lack of research on mental imagery relative to related topics (Pearson, 2014). Nevertheless, numerous findings support the notion of shared neural circuits between perception and imagery (for reviews see Kosslyn et al., 2006; Pearson, Naselaris, Holmes, and Kosslyn, 2015), accompanying the findings about top-down influences mentioned above.

1.6 OUTLINE OF THIS THESIS

This thesis sets out to investigate the functional architecture of several cortical areas in the human visual system during visual perception as well as imagery. It makes use of ultra-high field fMRI that allows us to map and decode brain activation at an unprecedented spatial resolution. Top-down influences on the visual system that are employed during mental visual imagery are investigated.

Chapter 2 (“Overlapping feature representations in human visual cortex: Binocular disparity and axis-of-motion tuning mapped in human area MT using ultra-high field fMRI”) is focused on depth perception and understanding the functional architecture of a higher-level visual area (human area MT or hMT) with respect to the computational properties and encoding of features. A regular, mesoscopic preference organisation for binocular disparity (visual depth) is shown. Moreover, this depth tuning is accompanied by

an additional tuning for direction-of-motion. This dual tuning is expected from animal models but shown for the first time in human area MT demonstrating a fine-grained organisation of cortical areas that code for multiple visual features simultaneously.

In [Chapter 3](#) (“Decoding the direction of imagined visual motion using 7 T ultra-high field fMRI”) a functional organisation similar to the one revealed in [Chapter 2](#) is observed, when subjects not only perceive a visual stimulus, but strikingly when they only imagine it. In this experiment, the direction of imagined visual motion can be decoded even from primary visual areas with remarkable decoding accuracies in single subjects. Top-down influences during imagery activate the early visual areas in a remarkably similar way to activations during perception.

Finally, in [Chapter 4](#) (“Reconstructing and decoding imagined letters from the visual system using ultra-high field fMRI”), a further and more precise form of visual imagery is decoded from areas across the visual system. Subjects imagine one of four letters and data recorded from early visual areas is used to reconstruct the imagined letter in an image-like form.

A discussion of these findings is presented in [Chapter 5](#). In summary, this thesis provides new evidence on the fine-grained organisation of visual areas in the human brain during perception and during imagery while employing state-of-the-art ultra-high resolution functional imaging techniques.

REFERENCES

- Albright, T. D. (1984, December). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of neurophysiology*, 52(6), 1106–30. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6520628>
- Backus, B. T., Fleet, D. J., Parker, A. J., & Heeger, D. J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of neurophysiology*, 86(4), 2054–2068.
- Bandettini, P. A. (2012). Functional MRI: A confluence of fortunate circumstances. *NeuroImage*, 61, 3–11. doi:10.1016/j.neuroimage.2012.01.130
- Cheng, K., Waggoner, R. A., & Tanaka, K. (2001, October). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron*, 32(2), 359–74. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11684004>
- De Martino, F., Zimmermann, J., Muckli, L., Uğurbil, K., Yacoub, E., & Goebel, R. (2013, January). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. *PloS one*, 8(3), e60514. doi:10.1371/journal.pone.0060514
- Epstein, R. & Kanwisher, N. (1998, April). A cortical representation of the local visual environment. *Nature*, 392(6676), 598–601. doi:10.1038/33402
- Feinberg, D. A., Harel, N., Ramanna, S., Uğurbil, K., & Yacoub, E. (2008). Submillimeter Single-shot 3D GRASE with Inner Volume Selection for T2 weighted fMRI applications at 7 Tesla 1. *Proc. Intl. Soc. Mag. Reson. Med.* 16, 37(4), 2373.
- Feinberg, D. A. & Yacoub, E. (2012). The rapid development of high speed, resolution and precision in fMRI. *NeuroImage*, 62(2), 720–725. doi:10.1016/j.neuroimage.2012.01.049. arXiv: NIHMS150003
- Felleman, D. J. & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, N.Y. : 1991)*, 1(1), 1–47. doi:10.1093/cercor/1.1.1
- Ferguson, E. S. (1977, August). The Mind's Eye: Nonverbal Thought in Technology. *Science (New York, N.Y.)* 197(4306), 827–36. doi:10.1126/science.197.4306.827
- Gilbert, C. D. & Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. *Neuron*, 54(5), 677–696. doi:10.1016/j.neuron.2007.05.019

- Goebel, R., Muckli, L., & Kim, D. S. (2011). Visual System. In J. K. Mai & G. Paxinos (Eds.), *The human nervous system* (3rd, pp. 1301–1327). Academic Press. Retrieved from <https://books.google.nl/books?id=J4cDpsl2rrQC>
- Goncalves, N. R., Ban, H., Sanchez-Panchuelo, R. M., Francis, S. T., Schluppeck, D., & Welchman, A. E. (2015). 7 Tesla fMRI Reveals Systematic Functional Organization for Binocular Disparity in Dorsal Visual Cortex. *Journal of Neuroscience*, 35(7), 3056–3072. doi:10.1523/JNEUROSCI.3047-14.2015
- Gregg, V., Winer, G. A., Cottrell, J. E., Hedman, K. E., & Fournier, J. S. (2001). The persistence of a misconception about vision after educational interventions. *Psychonomic Bulletin & Review*, 8(3), 622–6. doi:10.3758/BF03196199
- Harel, N. (2012). Ultra high resolution fMRI at ultra-high field. *NeuroImage*, 62(2), 1024–1028. doi:10.1016/j.neuroimage.2012.01.018
- Hirsch, S., Reichold, J., Schneider, M., Székely, G., & Weber, B. (2012). Topology and hemodynamics of the cortical cerebrovascular system. *Journal of Cerebral Blood Flow & Metabolism*, 32(6), 952–967. doi:10.1038/jcbfm.2012.39
- Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, 148, 574–591. doi:10.1113/jphysiol.2009.174151
- Hubel, D. H. & Wiesel, T. N. (1962, January). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160, 106–54.
- Hubel, D. H. & Wiesel, T. N. (1968, January). Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 67, 40–113. doi:10.1113/jphysiol.1968.sp008455
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997, June). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 17(11), 4302–11. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9151747>
- Kosslyn, S. M. (1980). *Image and Mind*. Harvard University Press.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001, September). Neural foundations of imagery. *Nature reviews. Neuroscience*, 2(9), 635–42. doi:10.1038/35090055
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006, March). *The Case for Mental Imagery*. Oxford University Press. doi:10.1093/acprof:oso/9780195179088.001.0001

- Kourtzi, Z. & Kanwisher, N. (2001, August). Representation of perceived object shape by the human lateral occipital complex. *Science (New York, N.Y.)* 293(5534), 1506–9. doi:[10.1126/science.1061133](https://doi.org/10.1126/science.1061133)
- Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002, March). Revising the Visualizer-Verbalizer Dimension: Evidence for Two Types of Visualizers. *Cognition and Instruction*, 20(1), 47–77. doi:[10.1207/S1532690XCI2001_3](https://doi.org/10.1207/S1532690XCI2001_3)
- Kozhevnikov, M., Kosslyn, S. M., & Shephard, J. (2005, June). Spatial versus object visualizers: a new characterization of visual cognitive style. *Memory & cognition*, 33(4), 710–26. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16248335>
- Kozhevnikov, M., Kozhevnikov, M., Yu, C. J., & Blazhenkova, O. (2013, June). Creativity, visualization abilities, and visual cognitive style. *The British journal of educational psychology*, 83(Pt 2), 196–209. doi:[10.1111/bjep.12013](https://doi.org/10.1111/bjep.12013)
- Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., . . . Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, 89(12), 5675–9. doi:[10.1073/pnas.89.12.5675](https://doi.org/10.1073/pnas.89.12.5675)
- McKeefry, D. J. & Zeki, S. M. (1997). The position and topography of the human colour centre as revealed by functional magnetic resonance imaging. *Brain*, 120(12), 2229–2242. doi:[10.1093/brain/120.12.2229](https://doi.org/10.1093/brain/120.12.2229)
- Miller, A. I. (1996). *Insights of Genius*. New York, NY: Springer New York. doi:[10.1007/978-1-4612-2388-7](https://doi.org/10.1007/978-1-4612-2388-7)
- Motes, M. A., Malach, R., & Kozhevnikov, M. (2008, November). Object-processing neural efficiency differentiates object from spatial visualizers. *Neuroreport*, 19(17), 1727–31. doi:[10.1097/WNR.0b013e328317f3e2](https://doi.org/10.1097/WNR.0b013e328317f3e2)
- Mountcastle, V. B. (1957, July). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of neurophysiology*, 20(4), 408–34. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/13439410>
- Mountcastle, V. B. (1997, April). The columnar organization of the neocortex. *Brain*, 120(4), 701–722. doi:[10.1093/brain/120.4.701](https://doi.org/10.1093/brain/120.4.701)
- Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Uğurbil, K., . . . Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology*, 25(20), 2690–2695. doi:[10.1016/j.cub.2015.08.057](https://doi.org/10.1016/j.cub.2015.08.057)

- Ogawa, S. & Lee, T. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the ...* 87(24), 9868–72. doi:[10.1073/pnas.87.24.9868](https://doi.org/10.1073/pnas.87.24.9868)
- Pearson, J. (2014). New Directions in Mental-Imagery Research: The Binocular-Rivalry Technique and Decoding fMRI Patterns. *Current Directions in Psychological Science*, 23, 178–183. doi:[10.1177/0963721414532287](https://doi.org/10.1177/0963721414532287)
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015, October). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, 19(10), 590–602. doi:[10.1016/j.tics.2015.08.003](https://doi.org/10.1016/j.tics.2015.08.003)
- Poltrock, S. & Brown, P. (1984). Individual differences in visual imagery and spatial ability. *Intelligence*, 138, 93–138. doi:[10.1016/0160-2896\(84\)90019-9](https://doi.org/10.1016/0160-2896(84)90019-9)
- Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 80(1), 1–24. doi:[10.1037/h0034650](https://doi.org/10.1037/h0034650)
- Rorden, C. & Karnath, H.-O. (2004). Using human brain lesions to infer function: a relic from a past era in the fMRI age? *Nature reviews. Neuroscience*, 5(10), 813–9. doi:[10.1038/nrn1521](https://doi.org/10.1038/nrn1521)
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., ... Tootell, R. B. H. (1995, May). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science (New York, N.Y.)* 268(5212), 889–93. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7754376>
- Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N. K., & Ugurbil, K. (2007, April). Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. *NeuroImage*, 35(2), 539–52. doi:[10.1016/j.neuroimage.2006.12.030](https://doi.org/10.1016/j.neuroimage.2006.12.030)
- Smith, A. M. (2014). *From Sight to Light: The Passage from Ancient to Modern Optics*. University of Chicago Press.
- Smith, F. W. & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences of the United States of America*, 107(46), 20099–103. doi:[10.1073/pnas.1000233107](https://doi.org/10.1073/pnas.1000233107). arXiv: [arXiv:1408.1149](https://arxiv.org/abs/1408.1149)
- Tootell, R. B. H., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., & Dale, A. M. (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 95(3), 811–817. doi:[10.1073/pnas.95.3.811](https://doi.org/10.1073/pnas.95.3.811)
- Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., ... Belliveau, J. W. (1995, April). Functional analysis of human MT and re-

- lated visual cortical areas using magnetic resonance imaging. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 15(4), 3215–30. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7722658>
- Tye, M. (2000). *The Imagery Debate*. London: The MIT Press.
- Uğurbil, K., Adriany, G., Andersen, P., Chen, W., Garwood, M., Gruetter, R., ... Zhu, X.-H. (2003, December). Ultrahigh field magnetic resonance imaging and spectroscopy. *Magnetic Resonance Imaging*, 21(10), 1263–1281. doi:[10.1016/j.mri.2003.08.027](https://doi.org/10.1016/j.mri.2003.08.027)
- Uludağ, K., Müller-Bierl, B., & Uğurbil, K. (2009, October). An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *NeuroImage*, 48(1), 150–65. doi:[10.1016/j.neuroimage.2009.05.051](https://doi.org/10.1016/j.neuroimage.2009.05.051)
- Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science (New York, N.Y.)* 255(5043), 419–423. doi:[10.1126/science.1734518](https://doi.org/10.1126/science.1734518)
- Vetter, P., Smith, F. W., & Muckli, L. (2014, June). Decoding sound and imagery content in early visual cortex. *Current biology : CB*, 24(11), 1256–62. doi:[10.1016/j.cub.2014.04.020](https://doi.org/10.1016/j.cub.2014.04.020)
- Weiner, K. S. & Zilles, K. (2015). The anatomical and functional specialization of the fusiform gyrus. *Neuropsychologia*, 1–15. doi:[10.1016/j.neuropsychologia.2015.06.033](https://doi.org/10.1016/j.neuropsychologia.2015.06.033)
- Winer, G. A., Cottrell, J. E., Gregg, V., Fournier, J. S., & Bica, L. A. (2002). Fundamentally misunderstanding visual perception: Adults' belief in visual emissions. *American Psychologist*, 57(6-7), 417–424. doi:[10.1037/0003-066X.57.6-7.417](https://doi.org/10.1037/0003-066X.57.6-7.417)
- Wraga, M., Thompson, W. L., Alpert, N. M., & Kosslyn, S. M. (2003, July). Implicit transfer of motor strategies in mental rotation. *Brain and Cognition*, 52(2), 135–143. doi:[10.1016/S0278-2626\(03\)00033-2](https://doi.org/10.1016/S0278-2626(03)00033-2)
- Yacoub, E., Harel, N., & Uğurbil, K. (2008, July). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(30), 10607–12. doi:[10.1073/pnas.0804110105](https://doi.org/10.1073/pnas.0804110105)
- Yacoub, E., Shmuel, A., Logothetis, N. K., & Uğurbil, K. (2007, October). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *NeuroImage*, 37(4), 1161–77. doi:[10.1016/j.neuroimage.2007.05.020](https://doi.org/10.1016/j.neuroimage.2007.05.020)

- Yacoub, E., Shmuel, A., Pfeuffer, J., Van de Moortele, P.-F., Adriany, G., Andersen, P., ... Hu, X. (2001). Imaging brain function in humans at 7 Tesla. *Magnetic Resonance in Medicine*, 45(4), 588–94. doi:[10.1002/mrm.1080](https://doi.org/10.1002/mrm.1080)
- Young, T. (1802). On the Theory of Light and Colours. *Phil. Trans. R. Soc. Lond.* 92(January), 12–48. doi:[10.1098/rstl.1802.0004](https://doi.org/10.1098/rstl.1802.0004)
- Zilles, K. & Amunts, K. (2010, February). Centenary of Brodmann's map — conception and fate. *Nature Reviews Neuroscience*, 11(2), 139–145. doi:[10.1038/nrn2776](https://doi.org/10.1038/nrn2776)
- Zimmermann, J., Goebel, R., De Martino, F., Van de Moortele, P.-F., Feinberg, D. A., Adriany, G., ... Yacoub, E. (2011, December). Mapping the Organization of Axis of Motion Selective Features in Human Area MT Using High-Field fMRI. *PLoS ONE*, 6(12), e28716. doi:[10.1371/journal.pone.0028716](https://doi.org/10.1371/journal.pone.0028716)

Chapter 2

Overlapping feature representations in human visual cortex: Binocular disparity and axis-of-motion tuning mapped in human area MT

based on: Emmerling, T. C., Valente, G., Zimmermann, J., Kemper, V. G., Frost, M., & Goebel, R. (*submitted*). Overlapping feature representations in human visual cortex: Binocular disparity and axis-of-motion tuning mapped in human area MT using ultra-high field fMRI.

We would like to thank Franziska Emmerling, Carmine Gnolo, Ömer Faruk Gülban, Roy Haast, and Dimo Ivanov for their help.

ABSTRACT

The ability to vividly perceive depth is strongly dependent on binocular disparity. While DeAngelis and Newsome (1999) demonstrated the importance of middle temporal area (MT) for stereoscopic depth perception in the macaque, corresponding findings for homologue human middle temporal area (hMT) have not been reported yet, mainly due to limits in spatial resolution of current neuroimaging techniques. We studied disparity tuning in hMT using submillimetre functional resolution ultra-high field (7 Tesla) functional magnetic resonance imaging (fMRI) and found a systematic fine-grained tuning organisation. At the same time, we could replicate axis-of-motion tunings in hMT as reported in Zimmermann et al. (2011). We, thus, show for the first time overlapping tuning maps for two distinct visual features, i. e. binocular disparity and direction-of-motion, in an extrastriate area of human visual cortex using a fully crossed experimental design to shed light on how the brain maps multiple basic visual features in the same cortical area.

2.1 INTRODUCTION

In recent years, it has become feasible to map the organisation within cortical areas of the human brain *in vivo* at a mesoscopic level using ultra-high field (UHF) fMRI (Cheng, Waggoner, & Tanaka, 2001; Goncalves et al., 2015; Yacoub, Harel, & Ugurbil, 2008; Yacoub, Shmuel, Logothetis, & Ugurbil, 2007; Zimmermann et al., 2011). A characterisation of the functional organisation of local neural circuits in the mammal cortex was already possible in animal models and led to the influential concept of a (hyper-)columnar organisation in the cortex (Hubel & Wiesel, 1974). The advent of submillimetre resolution UHF fMRI allows the replication and validation of the findings from cat and macaque cortex in the human. In this study we focus on the prominent mid-level area of human visual cortex – human middle temporal area (hMT) – and demonstrate two orthogonal but overlapping systematic mesoscopic organisations for the representations of binocular disparity and axis-of-motion.

Binocular disparity is an important cue in human vision enabling the vivid perception of depth in everyday life (Qian, 1997). Through a small relative shift between the projections of one point in visual space onto the retinas of the two eyes, humans, as well as many other binocular species such as monkeys, can infer the relative distance of objects. The computations that are necessary to fuse the images from the two retinas and extract the binocular disparity take place in areas along the visual hierarchy (Gonzalez & Perez, 1998). Optical imaging and electrophysiology studies in cat and macaque cortex showed groups of neurons that are selective for binocular disparity in V_1 (Grunewald & Skounbourdis, 2004), V_2 (Chen, Lu, & Roe, 2008; Hubel & Livingstone, 1987; Kara & Boyd, 2009; Roe, 1995; Ts’O, Roe, & Gilbert, 2001),

V₃/V_{3A} (Adams & Zeki, 2001; Anzai, Chowdhury, & DeAngelis, 2011; Hubel, Wiesel, Yeagle, Lafer-Sousa, & Conway, 2015), V₄ (Tanabe, Doi, Umeda, & Fujita, 2005), and MT/V₅ (DeAngelis, Cumming, & Newsome, 1998; DeAngelis & Newsome, 1999). Recently, studies using fMRI also demonstrated selectivity for disparity in the human in areas V₂, V₃, V_{3A} (Backus, Fleet, Parker, & Heeger, 2001; Nasr, Polimeni, & Tootell, 2016; Tsao et al., 2003), as well as V_{3B}/KO, V₇, and hMT+/V₅ (Preston, Li, Kourtzi, & Welchman, 2008). Goncalves et al., 2015 showed systematic disparity tuning maps in human V_{3A} using UHF fMRI. This is especially remarkable as it closely resembles the findings from macaque research (Adams & Zeki, 2001; Anzai et al., 2011; Hubel et al., 2015) and demonstrates a systematically organised mesoscopic representation of binocular disparity tuning in the human cortex.

A similar correspondence between macaque and human visual cortex should also exist in human area MT. In correspondence with findings from homologue macaque MT (Albright, 1984; Britten, Newsome, Shadlen, Movshon, & Celebrini, 1996; Diogo, Soares, Koulakov, Albright, & Gattass, 2003; Dubner & Zeki, 1971; Salzman, Murasugi, Britten, & Newsome, 1992), an axis-of-motion selective columnar organisation in hMT was shown by Zimmermann et al., 2011. As not only axis-of-motion tuning but also disparity tuning was repeatedly found in macaque MT/medial superior-temporal area (MST) (DeAngelis et al., 1998; DeAngelis & Newsome, 1999; DeAngelis & Uka, 2003; Roy, Komatsu, & Wurtz, 1992) an overlapping mesoscopic selectivity for these two features should also exist in hMT as proposed by DeAngelis and Newsome (1999, see also Figure 2.2a). However, a systematic disparity tuning in this area has not yet been shown.

In this study we reveal a systematic organisation of mesoscopic patches in hMT according to their tuning for different levels of

binocular disparity. We show, for the first time in humans, a disparity tuning in hMT that closely matches findings from macaque cortex (DeAngelis & Newsome, 1999). We also replicate earlier findings about an axis-of-motion tuning in hMT (Zimmermann et al., 2011), however we extend these findings and map both, disparity tuning and axis-of-motion tuning, in hMT at the same time with data from the same experiment. Finally, we compare axis-of-motion tunings across different levels of disparity and find differential effects for vertical and horizontal axes-of-motion, that might be related to findings from macaque cortex on a joint encoding of direction-of-motion and disparity (Smolyanskaya, Ruff, & Born, 2013).

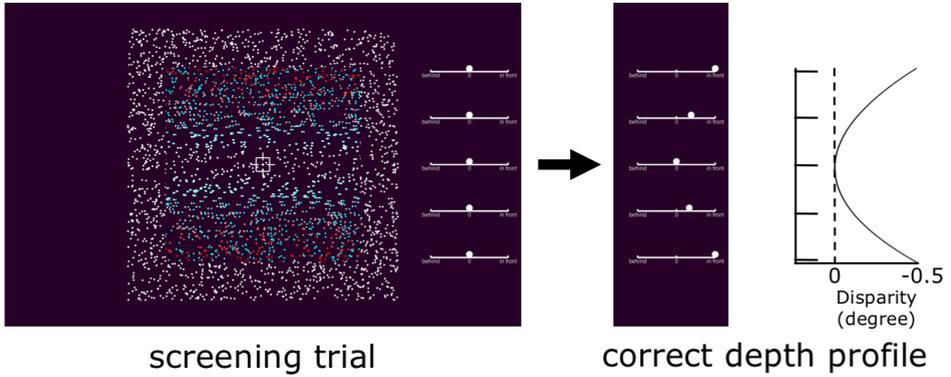
2.2 METHODS

SUBJECTS Ten healthy fMRI-experienced subjects (three females; age: 26.6 ± 3.3 years) with normal or corrected-to-normal vision volunteered for this study. Out of these subjects, seven subjects (two females; age: 26 ± 2.9 years) were selected to participate in an fMRI session based on their screening test scores (see below). They gave written informed consent and were paid for their participation. All procedures were conducted with approval from the local Ethical Committee of the Faculty of Psychology and Neuroscience at Maastricht University. All subjects were students or employees at Maastricht University.

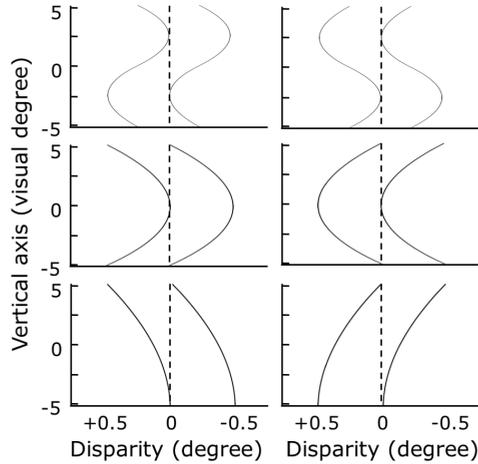
SCREENING SESSION All subjects attended a screening session to test their individual sensitivity for depth perception based on binocular disparity cues presented with anaglyph red-cyan glasses (026HT bright red and 116HT medium blue-g; LEE Filters, An-

dover, Hampshire, UK). The screening experiment was similar to the screening used by Georgieva, Peeters, Kolster, Todd, and Orban (2009). In each trial the subject was instructed to match a set of five visual analogue scales (VASs) to the perceived depth profile of a moving dots pattern at five sampling locations (marked by static dots in the moving dots pattern) using the computer mouse (see [Figure 2.1a](#)). The dot pattern was presented in a central square surrounded by flickering zero-disparity dots. The VAS were initially centred at the anchor '0' (disparity) and could be moved toward the anchors 'behind' (positive disparity) or 'in front' (negative disparity). The subject was instructed to fixate during the short 500 ms presentation of the dot pattern to prevent possible compensatory ocular vergence. Following a 1 s delay, the subject had the opportunity to display the stimulus again by pressing a button. The subject could trigger this repeated display for an unlimited number of times. When the subject finished the setting of all VAS the trial was completed with the press of another button. During the screening experiment dots moved along 12 depth patterns (see [Figure 2.1b](#)) that were presented at three different disparity amplitude ranges resulting in 36 different trials. The screening session began with a training phase in which the dot patterns were presented for an unlimited amount of time and performance scores were shown after each trial (correlation between VAS values and real depth profile converted to 0–100%). After the training phase subjects completed two sets of randomly presented trials (72 trials in total). A screening session lasted between 30 and 45 minutes.

EXPERIMENTAL TASK During the experiment subjects watched a pattern of moving and flickering dots through a pair of anaglyph red-cyan glasses similar to the stimulation in Gonçalves et al. (2015). The screen was filled with flickering white dots (20 Hz)



(a) Screening task



(b) Screening task depth patterns

Figure 2.1 Screening task. **(a)** A screening trial during the 500 ms presentation phase (left). The visual analogue scales (VASs) are in their original positions. The goal for the subject is to move the VAS sliders so that they reflect the perceived depth pattern of the square of dots on the screen that is presented with varying disparity. An exemplary setting of the VAS sliders is shown on the right (yielding a score of 99.7% correlation) together with the real depth profile of the shown pattern. **(b)** All depth patterns that were shown in the screening task at their maximal disparity range (0–0.5° visual degrees). Each pattern was shown three times with different disparity ranges.

providing a visual background at zero disparity. In four wedges around the fixation cross, dots were coherently moving in one of four directions (right, left, up, or down; speed of $8^\circ/\text{sec}$ and were presented at one of four disparities (visual angle of $+0.15^\circ$, $+0.05^\circ$, -0.05° , -0.15° ; far to near from observer; see [Figure 2.2b](#)). That is, each dot was presented as a pair of horizontally shifted red and cyan coloured dots. The red and cyan colours were calibrated to match the colour and luminance of the white (i. e. zero-disparity) dots when viewed through the filters. Therefore, the disparity dots appear to be placed at different levels of depth distance from the background of flickering white dots (and the viewer). The quadratic field of flickering white dots occupied a visual angle of 10° while the circle of wedges was shown at an eccentricity of 0.6° – 3.6° in radial direction. Wedges spanned 70° in polar angle and were spaced equally around fixation with gaps of 20° in polar angle between them. The visual stimulus was shown for 1000 ms followed by fixation cross for 500 ms. Eight repetitions of this stimulation formed one experimental block (12 s) during which the direction of motion and the disparity were kept constant. As in [Goncalves et al. \(2015\)](#) the polar angle at which the wedges were positioned was changed for each 1000 ms display interval to stimulate the entire 360° of polar angle. In each experimental run, 12 blocks for each disparity and direction-of-motion were presented in a fully crossed, pseudo-random order design resulting in 48 blocks per run.

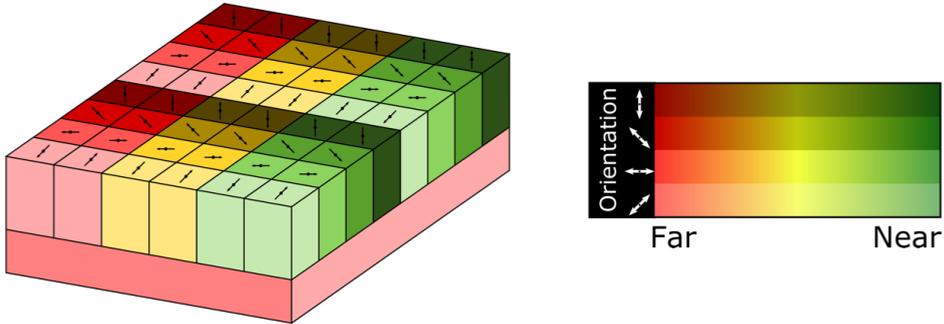
The fixation cross consisted of a white dot surrounded by a white square ($1^\circ \times 1^\circ$) and white nonius lines on the square. The subjects were instructed to maintain fixation during the entire experiment. In each 12 s block the horizontal or vertical nonius lines changed their relative position once to which the subjects were in-

structed to react by the press of a button. This task was included to ensure continuous fixation and attention.

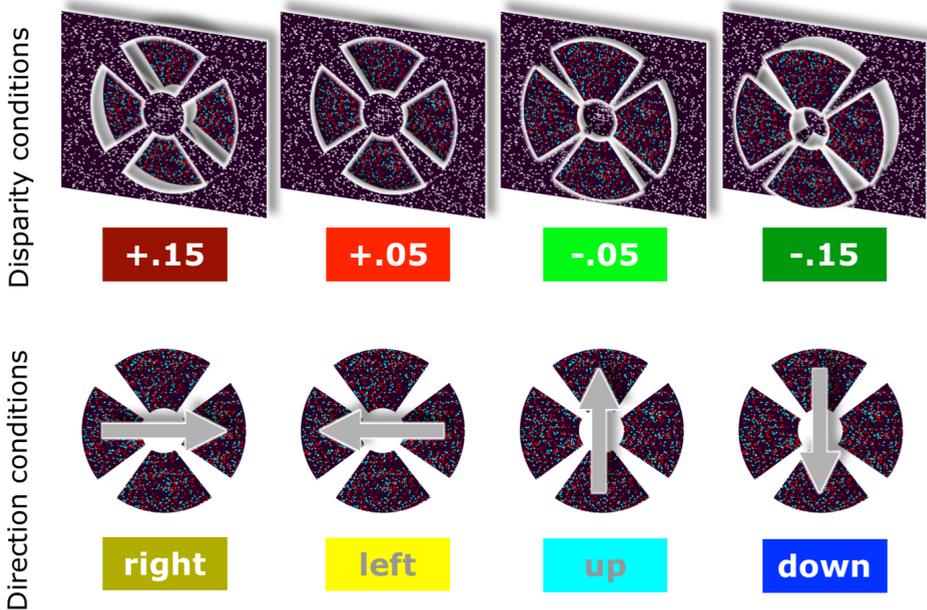
STIMULI Visual stimulation was created using PsychoPy (version 1.83.01; Peirce, 2007) and, in the scanning session, projected onto a frosted screen at the head end of the scanner bed using an LCD projector (Panasonic, No PT-EZ57OEL; Newark, NJ, USA). Behavioural responses to the attention task were collected through an MR compatible button box (Current Designs, 8-button response device, HHSC-2x4-C; Philadelphia, USA). Subjects wore an MR compatible eye frame equipped with colour filters for the right (026HT bright red; LEE Filters, Andover, Hampshire, UK) and left (116HT medium blue-green; LEE Filters, Andover, Hampshire, UK) eye. Colours of the visual stimulation were calibrated inside the scanner before the experiment so that there was no remaining cross-talk.

SCANNING SESSION At the beginning of the scanning session, we recorded an hMT+ localizer scan as in Emmerling, Zimmermann, Sorger, Frost, and Goebel (2016). Similar to Huk, Dougherty, and Heeger (2002), the stimulation included a field of dots moving in radial directions alternating with static dot fields. Dot fields were presented centrally and peripherally ($\pm 10^\circ$ of visual angle). After the localizer, four experimental runs were recorded.

MRI ACQUISITION Images were acquired with a Siemens Magnetom 7 T scanner (Siemens; Erlangen, Germany) and a 32-channel head-coil (Nova Medical Inc.; Wilmington, MA, USA). A T₁-weighted magnetisation prepared rapid acquisition gradient echo (3D-MP2RAGE; Marques et al., 2010) sequence (240 sagittal slices,



(a) Hypothesised idealised model of human area MT



(b) Experimental stimulation conditions

Figure 2.2 Hypothesised idealised model of human area MT and experimental stimulation conditions. **(a)** Similar to the model proposed by DeAngelis and Newsome (1999), this abstract organisation of human area MT shows mappings for direction tuning as well as for disparity tuning. **(b)** The eight experimental stimulation conditions are visualised. While the background field of white flickering dots was the same in every condition, an array of four polar wedges around the fixation cross was shown at four different disparities (upper row; visual angle of $+15^\circ$: very far from observer, $+5^\circ$: far from observer, -5° : near to observer, -15° : very near to observer; presented using anaglyph colours) and with dots moving in one of four directions (lower row; right, left, up, down).

matrix = 320×320 , voxel size = $0.7 \times 0.7 \times 0.7 \text{ mm}^3$, first inversion time $TI_1 = 900 \text{ ms}$, second inversion time $TI_2 = 2750 \text{ ms}$, echo time (TE) = 2.46 ms , repetition time (TR) = 5000 ms , first nominal flip angle = 5° , second nominal flip angle = 3°) was used to acquire an anatomical dataset. Very high-resolution functional images were recorded using gradient echo (T_2^* weighted) echo-planar imaging (EPI; Moeller et al., 2010) with the following parameters: TE = 26 ms , TR = 3000 ms , generalised auto-calibrating partially parallel acquisitions (GRAPPA) factor = 3, multi-band factor = 2, nominal flip angle = 55° , number of slices = 82, matrix = 186×186 , voxel size = $0.8 \times 0.8 \times 0.8 \text{ mm}^3$. The field-of-view included occipital, temporal, and parietal areas but did not cover large parts of frontal cortex (see Figure 2.3a). Right after the anatomical data we acquired additional functional volumes (five volumes in the phase encoding direction of the experimental runs and five volumes with a reversed phase encoding direction) to correct for EPI distortions.

IMAGING DATA PREPROCESSING Functional and anatomical images were analysed using BrainVoyager QX (version 2.8; Brain Innovation; Maastricht, The Netherlands), custom code in MATLAB (version 2013a; The Mathworks Inc.; Natick, MA, USA), and PyMVPA (version 2.3; Hanke et al., 2009). Anatomical images were interpolated to a nominal voxel size of 0.8 mm isotropic to match the resolution of the functional data. Similar to Emmerling et al. (2016), the white/grey matter boundary was detected with the largely automatic segmentation tools of BrainVoyager QX. These tools employ an adaptive region-growing method that analyses intensity histograms and corrects topological errors in the detected borders before reconstructing the cortical surfaces (Goebel, Esposito, & Formisano, 2006; Kriegeskorte & Goebel, 2001).

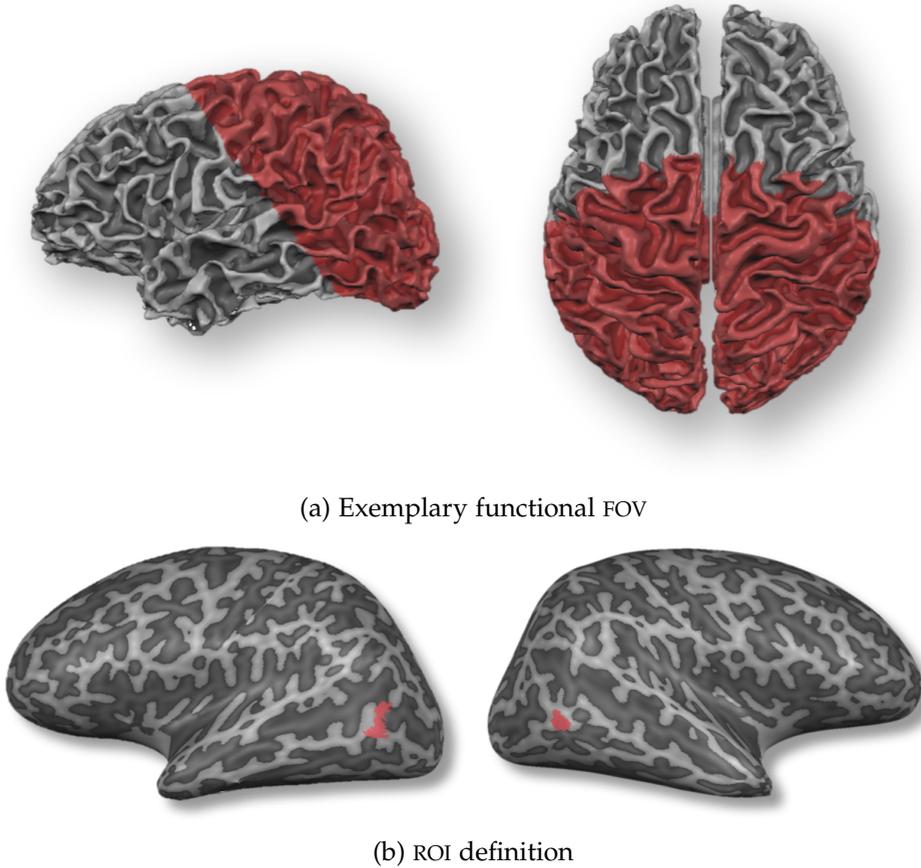


Figure 2.3 Exemplary functional FOV and region-of-interest (ROI) definition for subject 2. **(a)** The functional scans covered large parts of the visual cortex (red; shown on reconstructed cortical surface from left and from top) but excluded frontal and parietal areas. **(b)** Using data from an hMT+ localizer, human area MT (shown in red for left and right hemisphere) was defined on an inflated cortical surface by contrasting against MST. MT: middle temporal area; MST: middle superior temporal area.

In order to correct for distortions in the echo-planar imagings due to long echo-planar imaging (EPI) readout trains we recorded five functional volumes of normal and reversed phase encoding. These pairs of images are distorted in opposite directions and we used them to estimate the susceptibility-induced off-resonance field employing a method similar to that described in Andersson, Skare, and Ashburner (2003) as implemented in FSL (Smith et al., 2004). After performing 3D rigid body motion correction of the remaining functional runs (aligning all subsequent runs to the first functional run), the estimated off-resonance field was used to correct for EPI distortions. Furthermore, functional data were high-pass filtered using a general linear model (GLM) Fourier basis set of seven cycles sine/cosine per run (including linear trend removal). Functional runs were co-registered to the individual anatomical scan with an affine (9 parameter) transformation.

REGION-OF-INTEREST DEFINITION A human area MT region-of-interest (ROI) was defined using data from the hMT+ localizer scan projected onto an inflated surface reconstruction. In each hemisphere, area MT was distinguished from MST in the hMT+ complex (Emmerling et al., 2016; Huk et al., 2002) by thresholding the respective contrast at $q(\text{FDR}) < .001$ and restricting the resulting activation patch to 150 vertices (see Figure 2.3b). The surface patches were transformed back into volume space (from -1 mm to +3 mm from the grey/white matter segmentation boundary) to create the final volume ROIs.

SEARCHLIGHT ANALYSIS In each subject we mapped areas that allowed decoding the presented disparity or direction-of-motion by means of a searchlight analysis (Kriegeskorte, Goebel, & Bandettini, 2006). Similar to a previous study (Emmerling et al., 2016),

a sphere with a radius of four voxels was moved through the cortical ribbon (so that the spheres central voxel always lay within -1 mm to +3 mm from the grey/white matter segmentation border) and defined a feature set of 257 voxels. After z-scoring each run, the data were split into training and testing datasets in a leave-one-run-out cross-validation procedure. We extracted the average of 12 s (4 TRs), starting 6 s (2 TRs) after trial onset for each voxel and each trial of the spheres. These features were then used for a one-versus-one 4-class classification (predicted classes were chosen based on the maximum number of votes in all binary classifications) using a linear support vector machine (SVM; LIBSVM implementation in PyMVPA; Chang and Lin, 2011). Finally, we tested the resulting classification accuracies for significance by means of FDR-corrected Chi-square tests of the confusion matrix and projected the obtained thresholded p-value maps onto the inflated surface reconstruction.

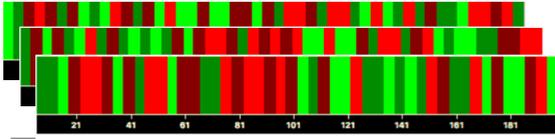
DISPARITY TUNING CURVES For each subject we assessed possible disparity tunings in area MT in a leave-one-run-out cross-validation procedure (see Figure 2.4). A GLM with predictors for all four disparity conditions was fit to data from three of the four runs (training data). After removing the mean t-value across voxels for each contrast (corresponding to one of the four disparities), these t-values were compared in each voxel. The mean removal was performed to eliminate signal biases towards near disparities as observed and corrected for in other studies (Adams & Zeki, 2001; DeAngelis & Uka, 2003; Tanabe et al., 2005). Each voxel was then labelled as preferring one of the four disparities based on the contrast with the highest t-value (see Figure 2.4, step 1). Subsequently, we fitted a GLM (same parameters as above) to the left-out run (testing data) and after mean removal (see above) the average

t-value was computed over all the voxels belonging to one class. This was done for all four groups of voxels labelled as preferring one of the four disparities (see [Figure 2.4](#), step 2) and repeated for four leave-one-run-out cross-validations (see [Figure 2.4](#), step 3). Finally, we computed a confusion matrix (CM) such that for each group of voxels the disparity condition with the maximum average t-value was counted over all cross-validations. If there was a perfect consistency of preferences between training and testing data this would lead to an identity matrix with '4's on the diagonal of the CM. That is: Voxels, that have the highest t-value for, e. g., disparity condition '-2' in the training data, also have the highest t-value for disparity condition '-2' in the testing data (and so forth for all classes and all cross-validation steps). We computed two accuracies for each CM (see [Figure 2.4](#), step 4):

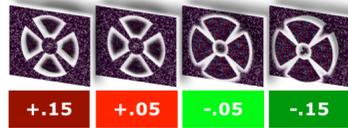
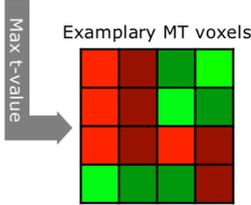
1. Two-class accuracy: sum of the values in the first and the fourth quadrant divided by sum of all values in the CM
2. Four-class accuracy: sum of the values on the main diagonal divided by sum of all values in the CM

PERMUTATION TESTING In order to test the computed accuracies (2-class and 4-class) for significance we performed 1000 permutations by scrambling the labels across conditions within each run, followed by the analysis pipeline described above. This gave us a distribution of accuracies under the null hypothesis to test the probability of the empirical accuracies in single subjects (see [Figure 2.4](#), step 5). Furthermore, we tested group-level significance by means of Wilcoxon Signed-Rank tests.

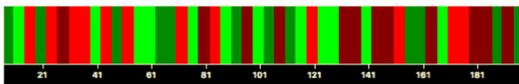
① Label MT voxels using training data



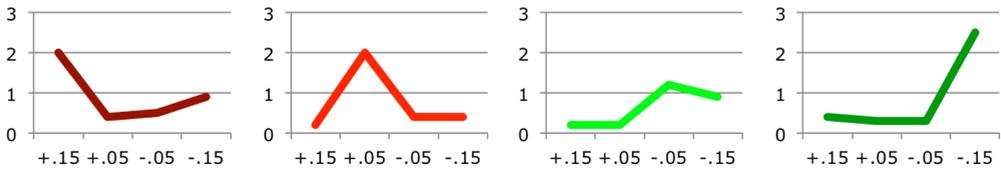
label voxels in MT ROI according to their disparity preference (maximal normalized t-value) in 3 of 4 runs (training data)



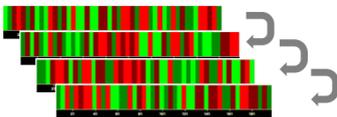
② Compute tuning curves for voxel group using testing data



compute average responses for all 4 labelled groups of voxels for all 4 disparities in the left-out run (testing data)



③ Repeat for 4 leave-one-run-out cross validations



④ Compute confusion matrix and accuracies

count maximum t-values for all disparities and groups of voxels over all cross validations

max response in testing data

	+.15	+.05	-.05	-.15
+.15	3	1	0	0
+.05	1	2	1	0
-.05	0	0	1	3
-.15	0	0	0	4

accuracy calculation:

$$ACC = \frac{TP + TN}{P + N}$$

2-class accuracy

	+.15	+.05	-.05	-.15
+.15	3	1	0	0
+.05	1	2	1	0
-.05	0	0	1	3
-.15	0	0	0	4

$$ACC_2 = \frac{7 + 8}{8 + 8} = .9375$$

4-class accuracy

	+.15	+.05	-.05	-.15
+.15	3	1	0	0
+.05	1	2	1	0
-.05	0	0	1	3
-.15	0	0	0	4

$$ACC_4 = \frac{5 + 5}{8 + 8} = .625$$

⑤ Repeat with within-run scrambled labels for permutation testing



Figure 2.4 Analysis pipeline to compute tuning curves for disparity in five steps (page 37). Please note that the analysis for direction tuning curves is equivalent to this pipeline. ACC: accuracy; TP: true positive; TN: true negative; P: all positive; N: all negative.

TUNING MAPS To visualise the organisation of disparity tunings we produced winner maps for the ROI assigning a colour to each voxel according to the most often occurring label over all four cross-validations. We then thresholded this map to only show voxels with a label that occurred in at least half of the cross-validations.

DIRECTION-OF-MOTION TUNING CURVES AND TUNING MAPS We performed the same analysis pipeline as outlined above (see also Figure 2.4) separately for the direction-of-motion conditions. This was possible because of the fully crossed design of the experiment. In the case of the direction-of-motion tuning curves the 2-class accuracy was computed for horizontal (left and right) versus vertical (up and down) directions-of-motion, i. e., axes-of-motion. We also produced tuning maps for the direction-of-motion tuning in the same way as for the disparity tuning.

DIRECTION-OF-MOTION TUNING MAPS AT DIFFERENT DISPARITIES We checked for possible differences between the direction-of-motion tuning in positive (i. e. far from the observer) versus negative (i. e. near to the observer) disparities. This was done separately for the horizontal and vertical axis-of-motion. We computed GLMs with predictors for all 16 conditions (4 disparities \times 4 directions) and calculated a preferred direction for every voxel in the MT ROI as outlined before (see Figure 2.4, step 1) for trials

with positive and negative disparities, thus, obtaining two different winner maps. Then, we applied a threshold to these winner maps so that only voxels with a label that occurred in at least half of the cross-validations were kept. Finally, we counted the number of voxels that had the same direction preference in both winner maps. We did that separately for horizontal and vertical direction preferences and standardised this number by the number of voxels with this direction preference in both maps (irrespective of similarity between maps). This way we obtained a consistency measure for horizontal and vertical direction preferences between positive and negative disparities.

To test the statistical significance of these consistencies we performed 10,000 permutation tests to obtain consistencies under the null hypothesis (there is no consistency between the two maps). Here, we scrambled the voxel labels in one of the two already thresholded maps and calculated the consistency as outlined above. A permutation test revealed significances for single subjects while we tested group-level significance by means of a Wilcoxon Signed-Ranks test of the found consistencies in each subject against the respective mean of permutations.

2.3 RESULTS

SCREENING DATA The average correlation indices between the depth profiles of the dot pattern and the setting of the visual analogue scales for 10 subjects are shown in [Figure 2.5](#). Subjects with an average correlation higher than $r = .75$ (S01—S07) were selected to participate in the scanning session.

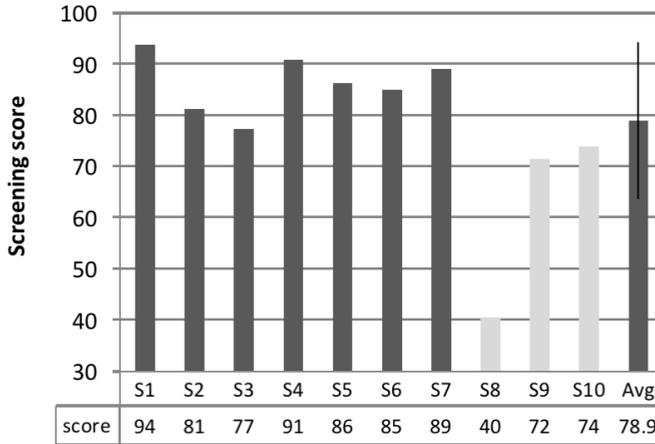
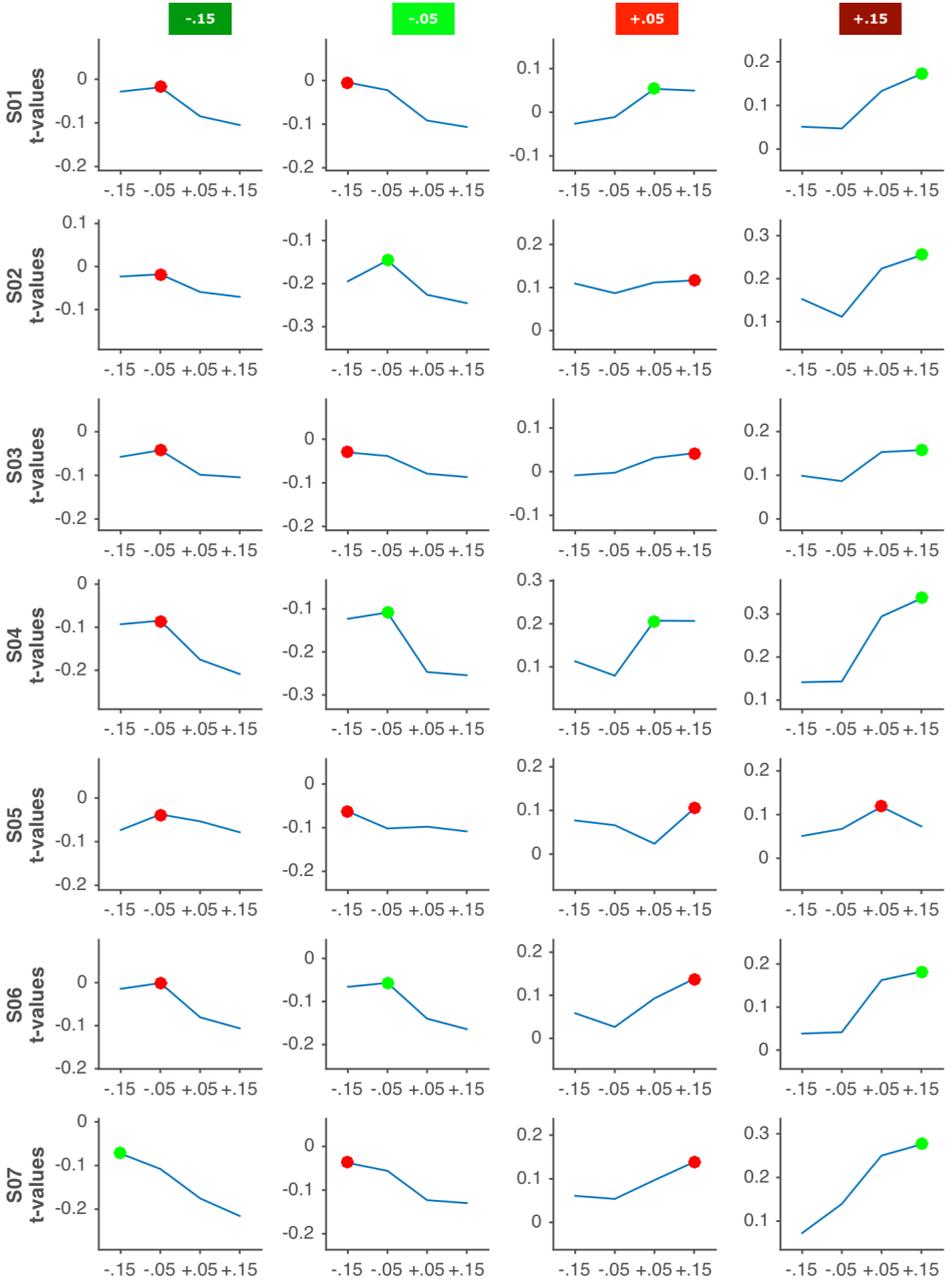


Figure 2.5 Behavioural results from screening session for 10 subjects. The screening score as an average correlation [%] between the setting of the visual analogue scales (VASs) and the actual depth pattern of the shown dot field across trials is shown for all subjects (S1 through S10). Subjects with an average correlation above 75 % were selected to participate in the scanning session (dark grey bars; subjects S1 through S7). The average correlation across all subjects is shown on the right (Avg). The error bar indicates one standard deviation.

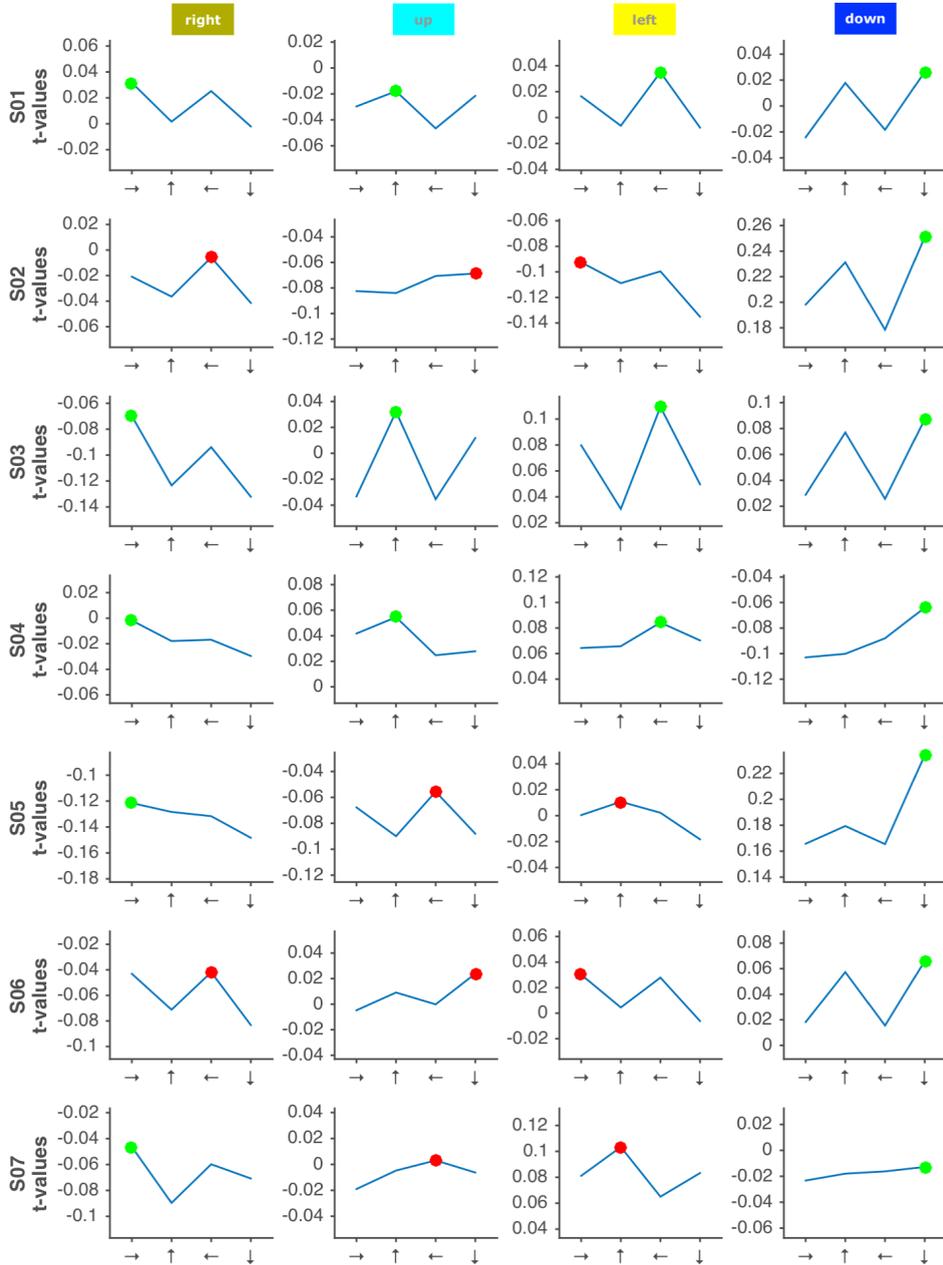
DISPARITY TUNING Disparity tuning curves for all subjects are shown in [Figure 2.6a](#). For the 2-class hypothesis, i. e. a tuning for positive (far) versus negative (near) disparities, we found significant tunings (as assessed by the accuracy of the confusion matrix) in all seven subjects that participated in the scanning session ($p < .05$; see [Figure 2.7a](#)). Group level statistics also showed a significant tuning ($W = 28$; $p = .016$). For the 4-class hypothesis, i. e. a tuning for all four levels of disparity, we found significant tunings in subjects S01 and S02 ($p < .05$). A group level statistical test also showed a significant tuning ($W = 26$; $p = .047$). Disparity tuning maps for subjects S01 and S03 are shown in [Figure 2.10](#).

DIRECTION-OF-MOTION TUNING Direction-of-motion tuning curves for all subjects are shown in [Figure 2.6b](#). For the 2-class hypothesis, i. e. a tuning for horizontal (left and right) versus vertical (up and down) axes-of-motion, we found significant tunings (as assessed by the accuracy of the confusion matrix) in subjects S01, S03, and S06 ($p < .05$; see [Figure 2.7b](#)). Group level statistics showed a significant tuning as well ($W = 28$; $p = .016$). For the 4-class hypothesis, i. e. a tuning for all four directions-of-motion, we found significant tunings in subjects S01 and S03 ($p < .05$). Furthermore, group level statistics showed a significant tuning ($W = 26$; $p = .047$). Direction-of-motion tuning maps for subjects S01 and S03 are shown in [Figure 2.10](#).

Figure 2.6 Tuning curves for disparity and direction-of-motion for subject S01 through S07 ([page 42](#) and [page 43](#)). **(a)** Disparity tuning curves for all four groups of voxels preferring a single disparity in the training data (subplots from left to right) are shown for all disparities in the test data (data points in each subplot; y-axis shows different disparities). **(b)** Direction-of-motion tuning curves for all four groups of voxels preferring a single direction-of-motion in the training data (subplots from left to right) are shown for all directions-of-motion in the test data (data points in each subplot; arrows indicate the direction-of-motion). **(a & b)** All t-values have their respective mean removed to eliminate large biases over classes. Dot markers in the line plots indicate the maximum in each tuning curve; green dots indicate maxima that match the preference in the training data, red dots indicate all other maxima.



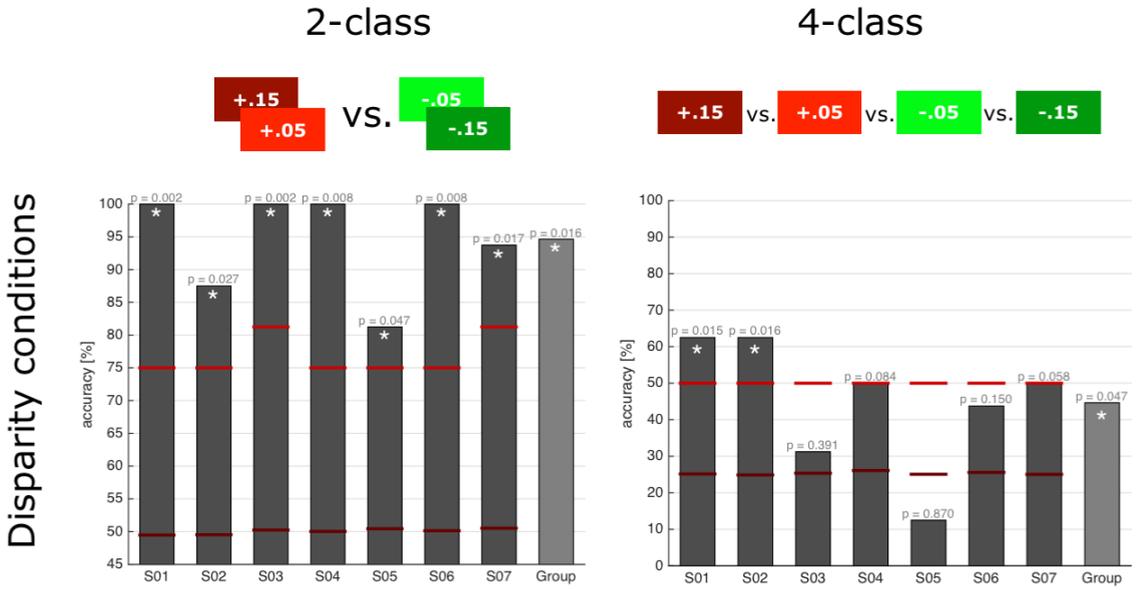
(a) Disparity tuning curves



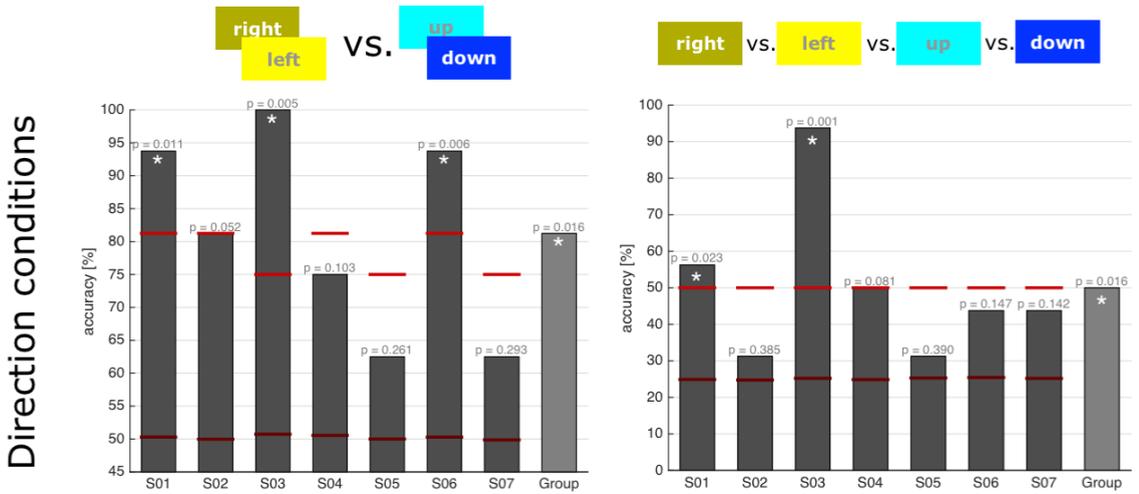
(b) Direction-of-motion tuning curves

SEARCHLIGHT RESULTS The decoding of the presented disparities and directions-of-motion by means of multi-voxel pattern analysis using a searchlight approach revealed significant cortical patches throughout the visual system in subjects So1, So2, So4, So6, and So7 for the disparity conditions (So4 and So6 show significant cortical patches in hMT) and no subject for the directions-of-motion conditions (see [Figure 2.8](#)).

Figure 2.7 Confusion matrix accuracies for disparity and direction-of-motion tunings for subject So1 through So7 ([page 45](#)). **(a)** Accuracies for a 2-class disparity tuning (left; positive versus negative disparity) and a 4-class disparity tuning (right; all four levels of disparity) in human area MT are shown. **(b)** Accuracies for a 2-class axes-of-motion tuning (left; horizontal versus vertical) and a 4-class direction-of-motion tuning (right; all four cardinal directions of motion) in human area MT are shown. **(a & b)** Dark red lines mark the single-subject mean of 1000 permutations; bright red bars mark the single-subject 95th percentile of 1000 permutations. Light grey bars labelled ‘Group’ show the respective average accuracy across all subjects. P-values are shown above the bars for each subject (assessed by testing against 1000 permutation) and for the group (assessed by Wilcoxon Signed-Rank tests against average permutation accuracies). Asterisks indicate significant accuracies ($p < .05$).



(a) Confusion matrix accuracies for disparity tunings



(b) Confusion matrix accuracies for direction-of-motion tunings

CONSISTENCY OF DIRECTION-OF-MOTION TUNING MAPS AT DIFFERENT DISPARITIES Results from the consistency analysis are shown in [Figure 2.9](#). On group level, the consistency values between the tuning maps at positive versus negative disparities revealed significant similarity for the vertical ($p = .031$; $W = 27$) but not for the horizontal axis-of-motion ($p = .156$; $W = 23$). For the horizontal axis-of-motion three subjects showed significant consistency while for the vertical axis-of-motion six subjects showed significant consistencies.

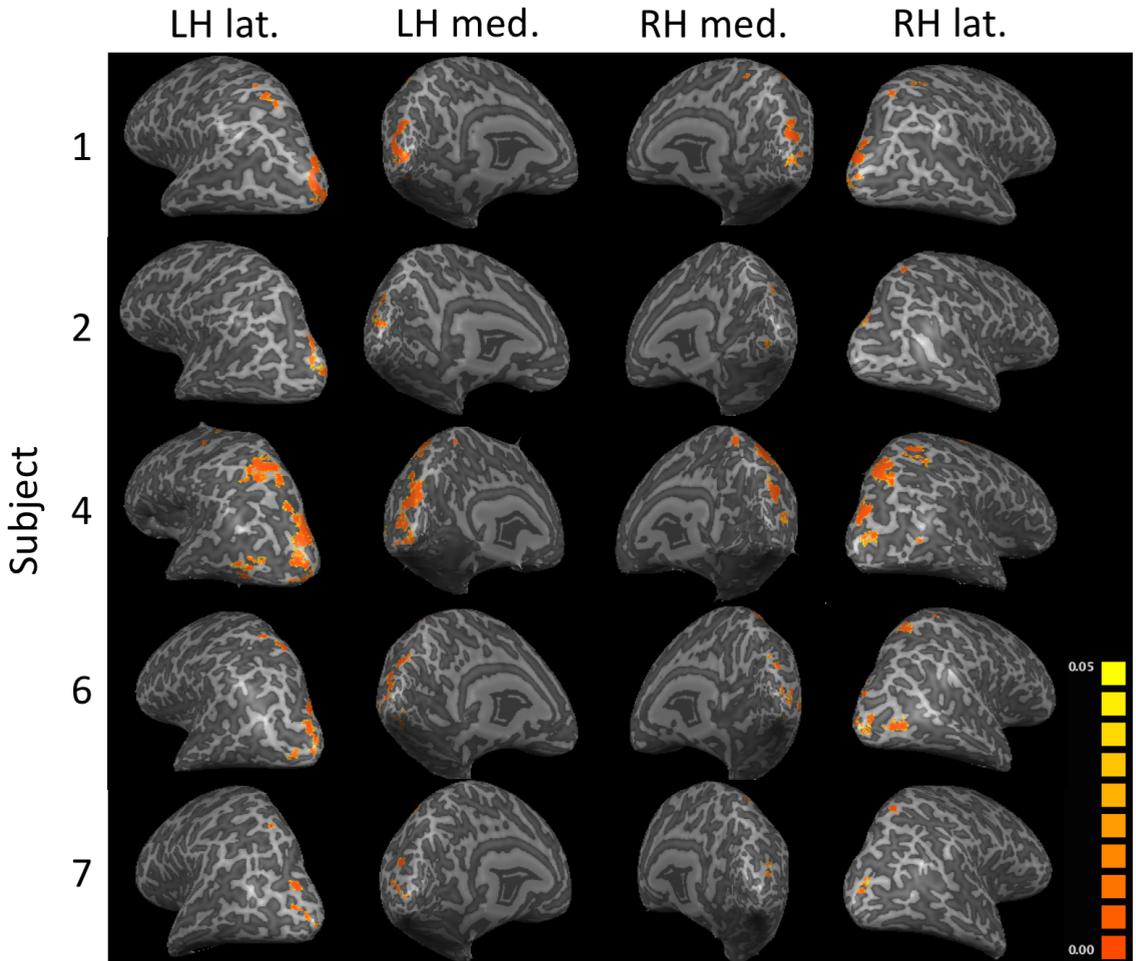


Figure 2.8 Searchlight maps for disparity decoding for five subjects. Highlighted patches in the left and right inflated hemispheres of subjects S01, S02, S04, S06, and S07 show significant decoding accuracies ($p < .05$ FDR corrected; 50 mm^2 cluster-threshold) for multi-voxel pattern analyses that run in spheres of voxels throughout the cortical ribbon of each subject. Warmer colours (orange > yellow) indicate lower p-value. In subjects S03 and S05 no significant patches were found that survived the cluster-thresholding. LH: left hemisphere; RH: right hemisphere; lat.: lateral view; med.: medial view.

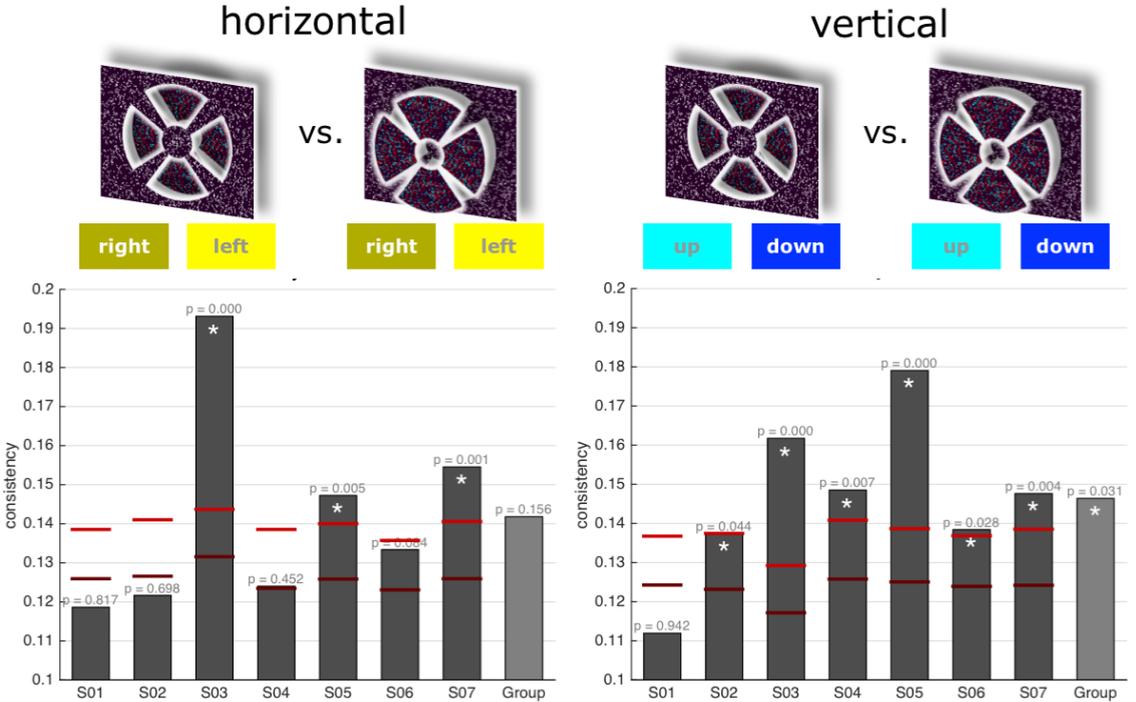
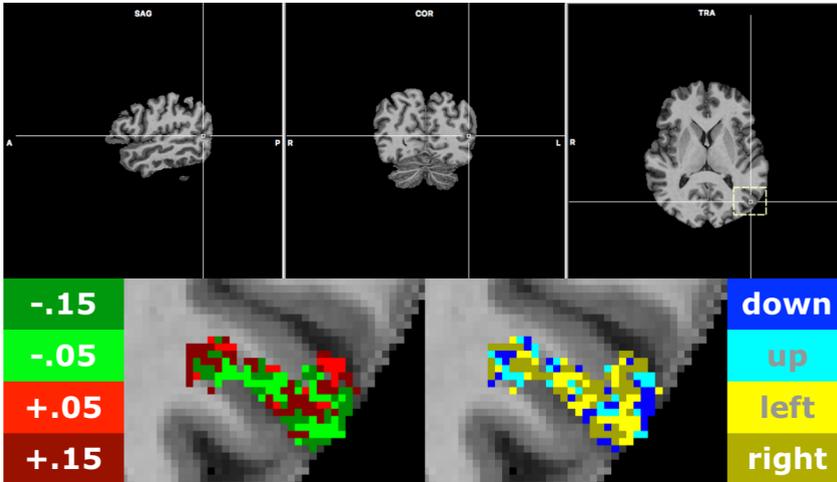
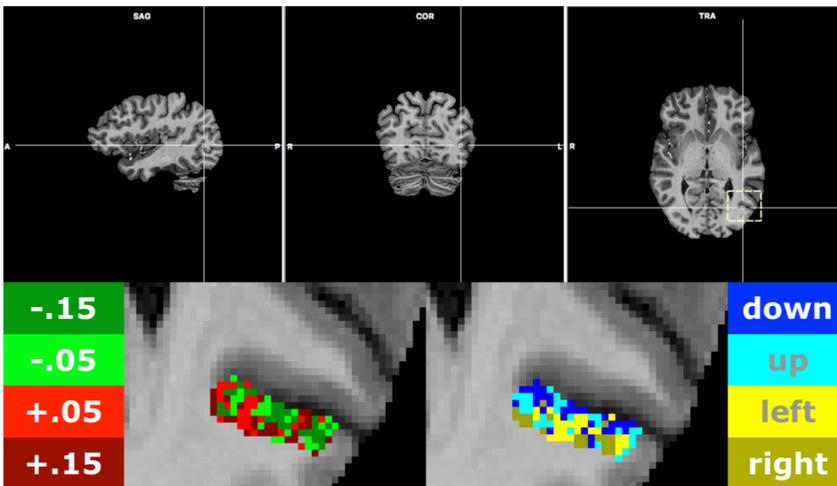


Figure 2.9 Consistencies between direction-of-motion tuning maps at different disparities. For subjects S01 through S07, consistencies are shown between horizontal (right and left) axis-of-motion tuning maps at positive (+.15; +.05) versus negative (-.05; -.15) disparities (left-hand side) and between vertical (up and down) axis-of-motion tuning maps at positive versus negative disparities (right-hand side). A high consistency indicates a similar tuning map in both groups of disparity conditions. That is, irrespectively of the disparity at which a moving dot pattern is presented the tuning map stays the same. Consistencies are standardised by the amount of voxels per class in each tuning map. Dark red lines mark the single-subject mean of 10,000 permutations; bright red bars mark the single-subject 95th percentile of 10,000 permutations. Light grey bars entitled 'Group' show the respective average consistency over all subjects. P-values are shown above the bars for each subject (assessed by testing against 10,000 permutations) and for the group (assessed by Wilcoxon Signed-Rank tests against average permutation accuracies). Asterisks indicate significant accuracies ($p < .05$).



(a) Tuning maps in hMT of subject S01



(b) Tuning maps in hMT of subject S03

Figure 2.10 Disparity tuning and direction-of-motion tuning maps in hMT. Winner maps for disparity tuning (left) and direction-of-motion tuning (right) are shown for subjects (a) S01 and (b) S03 in selected transversal slices in left-hemispheric hMT. Winner maps visualise the most often occurring label ('-.15', '-.05', '+.05', or '+.15' for disparity; 'down', 'up', 'left', or 'right' for direction-of-motion) per voxel. Colours are chosen to facilitate visual (2-class) pooling (i. e., 'near' versus 'far' for disparity; 'horizontal' versus 'vertical' for direction-of-motion). Maps are thresholded to only show voxels with a preference that was found in at least half of the cross-validations.

2.4 DISCUSSION

Discovering a tuning to more than one visual feature at the same time is of particular importance as it unravels the seemingly complicated multi-purpose architecture of visual areas. A decoding of visual features that are processed at various levels throughout the visual system allows for a better understanding of the whole computational pipeline of visual perception in the human brain.

This study set out to reveal mesoscopic dual feature representations of both binocular disparity and axis-of-motion tuning in hMT. Our results are the first to demonstrate hMT disparity tuning. Furthermore, we are able to replicate earlier findings (Zimmermann et al., 2011) of axis-of-motion tuning while unravelling the organisational structure of such a dual feature code. The overlap of systematic tunings for different visual features within one area was expected based on findings in homologue macaque area MT (DeAngelis & Newsome, 1999) and the tuning to both, binocular disparity and axis-of-motion, matches the macaque area MT model by DeAngelis and Newsome (1999, see also [Figure 2.2a](#)) closely.

We show significant disparity tunings ('near' versus 'far') in every single subjects as well as on a group level. We can even show significant 4-class disparity tunings in some subjects revealing a more fine-grained functional organisation of hMT that is also significant on a group level.

Beyond a significant tuning, however, we find a regular organisation observable for both, binocular disparity tuning as well as direction-of-motion tuning in single subjects (see [Figure 2.10](#)). A regular tuning map such as this is expected in light of similar results in macaque cortex (DeAngelis & Newsome, 1999).

Although this study was not primarily designed to examine a possible joint encoding of binocular disparity and direction-of-motion (Smolyanskaya et al., 2013) it provides first evidence that hMT represents both features in a largely separable manner. With our fully crossed experimental design it was possible to find consistent tunings for one feature (e.g. binocular disparity) despite the other feature changing (e.g. different directions-of-motion) by averaging across the irrelevant feature dimension Grunewald and Skoumbourdis, 2004; Qian and Andersen, 1997. If there was an interdependence in representations of disparity and direction-of-motion, i.e. a joint encoding, the proportion of neurons in hMT with such properties was so small that it did not change tuning accuracies for either separate feature domain significantly. This matches findings in macaque MT (Smolyanskaya et al., 2013) and suggests a multiplicative model of feature encoding in human area MT. Furthermore, it also provides evidence that our localisation of hMT in contrast to (human) MST was successful as a joint encoding for disparity and direction-of-motion has been repeatedly shown in macaque MST (Roy et al., 1992; Roy & Wurtz, 1990). If our hMT ROI definition had included MST it would have rendered cortical representations of the two features inseparable, assuming that this joint encoding is present in the human MST.

*Joint
encoding*

When we checked for differences between the direction-of-motion tuning maps at positive ('far') versus negative ('near') disparities we found an interaction with the axis-of-motion: For the vertical axis-of-motion (up and down) there was a significant consistency between the tuning maps at positive and negative disparities (as validated with permutation tests). However, for the horizontal axis-of-motion (right and left) this consistency was not significant. These systematic changes of direction-of-motion tunings across different levels of disparity could be explained by

possible joint encoding characteristics of subdivisions in hMT. Recently, Nadler et al. (2013) showed that in macaque MT 44% of the neurons are “opposite” cells preferring opposite depth-signs from motion parallax and binocular disparity (i. e. an “opposite” MT neuron would signal a near depth sign from motion parallax but a far depth sign from binocular disparity). While the exact functional role of opposite cells remains elusive, Nadler et al. (2013) hypothesise that these cells enable the detection of local mismatches between disparity and image motion and thereby allow the detection of moving objects in depth while the observer is in motion himself. As the computation of motion parallax information involves the integration of (extra-retinal cues and) horizontal axis-of-motion information (Nadler, Angelaki, & DeAngelis, 2008; Nadler, Nawrot, Angelaki, & DeAngelis, 2009) our results on direction-of-motion tuning consistency might directly relate to these opposite cells. We speculate that the hMT patches tuned for horizontal directions-of-motion might contain more opposite cells than the patches tuned for vertical directions-of-motion. Opposite cells could lead to a change in direction-of-motion preference across different levels of binocular disparity based on their opposite tuning for motion parallax on different disparity levels. This could explain the differential consistency effects we found in our data. An alternative explanation, however, would be that in some subjects the definition of the hMT ROI partially included MST and therefore included effects from a possible joint encoding present in MST (Roy et al., 1992; Roy & Wurtz, 1990). Follow-up experiments are needed to clarify the interaction of binocular disparity and direction-of-motion.

*Comparison
of tunings*

We found better searchlight accuracies for decoding binocular disparity than direction-of-motion. This might be a result from a ‘near-bias’ in visual areas that was very prominent in our data

and was shown in other studies (Adams & Zeki, 2001; DeAngelis & Uka, 2003; Tanabe et al., 2005). This near-bias might have led to enhanced performance of the searchlight analysis classifying between positive and negative disparities. Such a bias was probably not present for the different directions-of-motion. In combination with the limited volume of each searchlight sphere and the functional resolution we recorded at, this might explain the comparably worse performance of the searchlight decoding for directions-of-motion. For the searchlight decoding we did not test a pooled hypothesis of axes-of-motion, which might have led to better results. Especially the weak searchlight decoding results in hMT compare well to previous studies on direction-of-motion decoding (Kamitani & Tong, 2005).

Another factor that could have influenced the results on direction-of-motion decoding was the design of the visual stimulation. In this study we primarily wanted to map a disparity tuning in hMT and optimised our visual stimulation to enhance the subjective percept of visual depth. Therefore, we included a flickering dots background, a gapped stimulation pattern (blank periods of 500 ms), and manipulated binocular disparity and direction-of-motion only in four wedges around fixation to allow for a better perceptual contrast. In an experiment that primarily aimed for the mapping of direction-of-motion tuning, other experimental designs would have led to more pronounced results (Zimmermann et al., 2011).

We found disparity tunings in hMT using gradient echo fMRI sequences. While our sequence allowed for a field-of-view that included hMT as well as earlier visual areas we did not systematically check for disparity tunings in areas other than hMT. This was due to time constraints in the scanning session that prohibited the recording of a retinotopy to delineate these areas. It would be inter-

Conclusion

esting to record retinotopies in a follow-up experiment to compare our findings with results from Goncalves et al. (2015) and Nasr et al. (2016).

While our imaging resolution was sufficient to map two features at a mesoscopic scale in the same brain area using a gradient echo sequence, one could, in the future, also aim to map disparity and motion direction tunings at an even finer scale and with higher accuracies using spin-echo based fMRI sequences such as 3D GRASE (Feinberg, Harel, Ramanna, Uğurbil, & Yacoub, 2008) that have a higher specificity for local microvasculature compared to gradient echo sequences (De Martino et al., 2013; Uğurbil et al., 2003; Uluđag, Müller-Bierl, & Uğurbil, 2009; Yacoub et al., 2007). This would not only shed light on the fine-grained and possibly columnar architecture of hMT but could also address joint encoding properties of subpopulations within hMT. Furthermore, it will be interesting to apply the mapping approach we employed in this study to more complex visual paradigms in order to map and understand the functional architecture of human visual cognition beyond relatively simple features.

REFERENCES

- Adams, D. L. & Zeki, S. M. (2001). Functional organization of macaque V3 for stereoscopic depth. *Journal of neurophysiology*, 86(5), 2195–2203.
- Albright, T. D. (1984, December). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of neurophysiology*, 52(6), 1106–30. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6520628>
- Andersson, J. L. R., Skare, S., & Ashburner, J. (2003, October). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, 20(2), 870–88. doi:10.1016/S1053-8119(03)00336-7
- Anzai, A., Chowdhury, S. A., & DeAngelis, G. C. (2011). Coding of stereoscopic depth information in visual areas V3 and V3A. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(28), 10270–10282. doi:10.1523/JNEUROSCI.5956-10.2011. arXiv: NIHMS150003
- Backus, B. T., Fleet, D. J., Parker, A. J., & Heeger, D. J. (2001). Human cortical activity correlates with stereoscopic depth perception. *Journal of neurophysiology*, 86(4), 2054–2068.
- Britten, K. H., Newsome, W. T., Shadlen, M., Movshon, J. A., & Celebrini, S. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, 13(01), 87–100. doi:10.1017/S095252380000715X
- Chang, C.-C. & Lin, C.-J. (2011, April). Libsvm. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–27. doi:10.1145/1961189.1961199
- Chen, G., Lu, H. D., & Roe, A. W. (2008). A Map for Horizontal Disparity in Monkey V2. *Neuron*, 58(3), 442–450. doi:10.1016/j.neuron.2008.02.032
- Cheng, K., Waggoner, R. A., & Tanaka, K. (2001, October). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron*, 32(2), 359–74. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11684004>
- De Martino, F., Zimmermann, J., Muckli, L., Uğurbil, K., Yacoub, E., & Goebel, R. (2013, January). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. *PloS one*, 8(3), e60514. doi:10.1371/journal.pone.0060514

- DeAngelis, G. C., Cumming, B. G., & Newsome, W. T. (1998, August). Cortical area MT and the perception of stereoscopic depth. *Nature*, 394(6694), 677–80. doi:[10.1038/29299](https://doi.org/10.1038/29299)
- DeAngelis, G. C. & Newsome, W. T. (1999, February). Organization of disparity-selective neurons in macaque area MT. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 19(4), 1398–415. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9952417>
- DeAngelis, G. C. & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *Journal of neurophysiology*, 89(2), 1094–1111. doi:[10.1152/jn.00717.2002](https://doi.org/10.1152/jn.00717.2002)
- Diogo, A. C. M., Soares, J. G. M., Koulakov, A., Albright, T. D., & Gattass, R. (2003, May). Electrophysiological imaging of functional architecture in the cortical middle temporal visual area of Cebus apella monkey. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(9), 3881–98. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12736358>
- Dubner, R. & Zeki, S. M. (1971, December). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain research*, 35(2), 528–32. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/5002708>
- Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A., & Goebel, R. (2016, January). Decoding the direction of imagined visual motion using 7T ultra-high field fMRI. *NeuroImage*, 125, 61–73. doi:[10.1016/j.neuroimage.2015.10.022](https://doi.org/10.1016/j.neuroimage.2015.10.022)
- Feinberg, D. A., Harel, N., Ramanna, S., Uğurbil, K., & Yacoub, E. (2008). Sub-millimeter Single-shot 3D GRASE with Inner Volume Selection for T2 weighted fMRI applications at 7 Tesla 1. *Proc. Intl. Soc. Mag. Reson. Med.* 16, 37(4), 2373.
- Georgieva, S. S., Peeters, R., Kolster, H., Todd, J. T., & Orban, G. A. (2009, January). The processing of three-dimensional shape from disparity in the human brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(3), 727–42. doi:[10.1523/JNEUROSCI.4753-08.2009](https://doi.org/10.1523/JNEUROSCI.4753-08.2009)
- Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of Functional Image Analysis Contest (FIAC) data with BrainVoyager QX: From single-subject to cortically aligned group General Linear Model analysis and self-organizing group Independent Component Analysis. *Human Brain Mapping*, 27(5), 392–401. doi:[10.1002/hbm.20249](https://doi.org/10.1002/hbm.20249)

- Goncalves, N. R., Ban, H., Sanchez-Panchuelo, R. M., Francis, S. T., Schluppeck, D., & Welchman, A. E. (2015). 7 Tesla fMRI Reveals Systematic Functional Organization for Binocular Disparity in Dorsal Visual Cortex. *Journal of Neuroscience*, 35(7), 3056–3072. doi:[10.1523/JNEUROSCI.3047-14.2015](https://doi.org/10.1523/JNEUROSCI.3047-14.2015)
- Gonzalez, F. & Perez, R. (1998). Neural mechanisms underlying stereoscopic vision. *Progress in neurobiology*, 55(3), 191–224. doi:[10.1016/S0301-0082\(98\)00012-4](https://doi.org/10.1016/S0301-0082(98)00012-4)
- Grunewald, A. & Skoumbourdis, E. K. (2004, October). The Integration of Multiple Stimulus Features by V1 Neurons. *Journal of Neuroscience*, 24(41), 9185–9194. doi:[10.1523/JNEUROSCI.1884-04.2004](https://doi.org/10.1523/JNEUROSCI.1884-04.2004)
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009, January). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, 7(1), 37–53. doi:[10.1007/s12021-008-9041-y](https://doi.org/10.1007/s12021-008-9041-y)
- Hubel, D. H. & Livingstone, M. S. (1987). Segregation of form, color, and stereopsis in primate area 18. *The Journal of neuroscience*, 7(11), 3378–3415.
- Hubel, D. H. & Wiesel, T. N. (1974, December). Sequence regularity and geometry of orientation columns in the monkey striate cortex. *The Journal of comparative neurology*, 158(3), 267–93. doi:[10.1002/cne.901580304](https://doi.org/10.1002/cne.901580304)
- Hubel, D. H., Wiesel, T. N., Yeagle, E. M., Lafer-Sousa, R., & Conway, B. R. (2015). Binocular stereopsis in visual areas V-2, V-3, and V-3A of the macaque monkey. *Cerebral Cortex*, 25(4), 959–971. doi:[10.1093/cercor/bht288](https://doi.org/10.1093/cercor/bht288)
- Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002, August). Retinotopy and functional subdivision of human areas MT and MST. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(16), 7195–205. doi:[20026661](https://doi.org/20026661)
- Kamitani, Y. & Tong, F. (2005, May). Decoding the visual and subjective contents of the human brain. *Nature neuroscience*, 8(5), 679–85. doi:[10.1038/nn1444](https://doi.org/10.1038/nn1444)
- Kara, P. & Boyd, J. D. (2009). A micro-architecture for binocular disparity and ocular dominance in visual cortex. *Nature*, 458(7238), 627–631. doi:[10.1167/7.15.21](https://doi.org/10.1167/7.15.21)
- Kriegeskorte, N. & Goebel, R. (2001). An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical mr volumes. *NeuroImage*, 14(2), 329–346. doi:[10.1006/nimg.2001.0831](https://doi.org/10.1006/nimg.2001.0831)
- Kriegeskorte, N., Goebel, R., & Bandettini, P. A. (2006, March). Information-based functional brain mapping. *Proceedings of the National Academy of*

- Sciences of the United States of America*, 103(10), 3863–8. doi:[10.1073/pnas.0600244103](https://doi.org/10.1073/pnas.0600244103)
- Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.-F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage*, 49(2), 1271–1281. doi:[10.1016/j.neuroimage.2009.10.002](https://doi.org/10.1016/j.neuroimage.2009.10.002)
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E. J., Strupp, J. P., Harel, N., & Ugurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic Resonance in Medicine*, 63(5), 1144–1153. doi:[10.1002/mrm.22361](https://doi.org/10.1002/mrm.22361)
- Nadler, J. W., Angelaki, D. E., & DeAngelis, G. C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature*, 452(7187), 642–5. doi:[10.1038/nature06814](https://doi.org/10.1038/nature06814)
- Nadler, J. W., Barbash, D., Kim, H. R., Shimpi, S., Angelaki, D. E., & DeAngelis, G. C. (2013). Joint representation of depth from motion parallax and binocular disparity cues in macaque area MT. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(35), 14061–74, 14074a. doi:[10.1523/JNEUROSCI.0251-13.2013](https://doi.org/10.1523/JNEUROSCI.0251-13.2013)
- Nadler, J. W., Nawrot, M., Angelaki, D. E., & DeAngelis, G. C. (2009). MT Neurons Combine Visual Motion with a Smooth Eye Movement Signal to Code Depth-Sign from Motion Parallax. *Neuron*, 63(4), 523–532. doi:[10.1016/j.neuron.2009.07.029](https://doi.org/10.1016/j.neuron.2009.07.029)
- Nasr, S., Polimeni, J. R., & Tootell, R. B. H. (2016). Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. *Journal of Neuroscience*, 36(6), in press. doi:[10.1523/JNEUROSCI.3518-15.2016](https://doi.org/10.1523/JNEUROSCI.3518-15.2016)
- Peirce, J. W. (2007, May). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2), 8–13. doi:[10.1016/j.jneumeth.2006.11.017](https://doi.org/10.1016/j.jneumeth.2006.11.017)
- Preston, T. J., Li, S., Kourtzi, Z., & Welchman, A. E. (2008). Multivoxel pattern selectivity for perceptually relevant binocular disparities in the human brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(44), 11315–11327. doi:[10.1523/JNEUROSCI.2728-08.2008](https://doi.org/10.1523/JNEUROSCI.2728-08.2008)
- Qian, N. (1997). Binocular Disparity Review and the Perception of Depth. *Neuron*, 18(3), 359–368. doi:[10.1016/S0896-6273\(00\)81238-6](https://doi.org/10.1016/S0896-6273(00)81238-6)

- Qian, N. & Andersen, R. A. (1997). A physiological model for motion-stereo integration and a unified explanation of Pulfrich-like phenomena. *Vision Research*, 37(12), 1683–1698. doi:[10.1016/S0042-6989\(96\)00164-2](https://doi.org/10.1016/S0042-6989(96)00164-2)
- Roe, A. W. (1995). Visual topography in primate V2: multiple representation across functional stripes. *Zywnienie Czlowieka I Metabolizm*, 15(May).
- Roy, J.-P., Komatsu, H., & Wurtz, R. H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience*, 12(7), 2478–92. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1613542>
- Roy, J.-P. & Wurtz, R. H. (1990, November). The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. *Nature*, 348(6297), 160–162. doi:[10.1038/348160a0](https://doi.org/10.1038/348160a0)
- Salzman, D. C., Murasugi, C. M., Britten, K. H., & Newsome, W. T. (1992). Microstimulation in visual area MT: effects on direction discrimination performance. *The Journal of Neuroscience*, 12(6), 2331–2355.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., ... Matthews, P. M. (2004, January). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1, S208–19. doi:[10.1016/j.neuroimage.2004.07.051](https://doi.org/10.1016/j.neuroimage.2004.07.051)
- Smolyanskaya, A., Ruff, D. A., & Born, R. T. (2013). Joint tuning for direction of motion and binocular disparity in macaque MT is largely separable. *Journal of neurophysiology*, 110(12), 2806–16. doi:[10.1152/jn.00573.2013](https://doi.org/10.1152/jn.00573.2013)
- Tanabe, S., Doi, T., Umeda, K., & Fujita, I. (2005). Disparity-tuning characteristics of neuronal responses to dynamic random-dot stereograms in macaque visual area V4. *Journal of neurophysiology*, 94(4), 2683–2699. doi:[10.1152/jn.00319.2005](https://doi.org/10.1152/jn.00319.2005)
- Tsao, D. Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T. A., Mandeville, J. B., ... Tootell, R. B. H. (2003). Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron*, 39(3), 555–568. doi:[10.1016/S0896-6273\(03\)00459-8](https://doi.org/10.1016/S0896-6273(03)00459-8)
- Ts’O, D. Y., Roe, A. W., & Gilbert, C. D. (2001). A hierarchy of the functional organization for color, form and disparity in primate visual area V2. *Vision Research*, 41(10-11), 1333–1349. doi:[10.1016/S0042-6989\(01\)00076-1](https://doi.org/10.1016/S0042-6989(01)00076-1)
- Uğurbil, K., Adriany, G., Andersen, P., Chen, W., Garwood, M., Gruetter, R., ... Zhu, X.-H. (2003, December). Ultrahigh field magnetic resonance imaging and spectroscopy. *Magnetic Resonance Imaging*, 21(10), 1263–1281. doi:[10.1016/j.mri.2003.08.027](https://doi.org/10.1016/j.mri.2003.08.027)

- Uludağ, K., Müller-Bierl, B., & Uğurbil, K. (2009, October). An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *NeuroImage*, 48(1), 150–65. doi:[10.1016/j.neuroimage.2009.05.051](https://doi.org/10.1016/j.neuroimage.2009.05.051)
- Yacoub, E., Harel, N., & Uğurbil, K. (2008, July). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(30), 10607–12. doi:[10.1073/pnas.0804110105](https://doi.org/10.1073/pnas.0804110105)
- Yacoub, E., Shmuel, A., Logothetis, N. K., & Uğurbil, K. (2007, October). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *NeuroImage*, 37(4), 1161–77. doi:[10.1016/j.neuroimage.2007.05.020](https://doi.org/10.1016/j.neuroimage.2007.05.020)
- Zimmermann, J., Goebel, R., De Martino, F., Van de Moortele, P.-F., Feinberg, D. A., Adriany, G., . . . Yacoub, E. (2011, December). Mapping the Organization of Axis of Motion Selective Features in Human Area MT Using High-Field fMRI. *PLoS ONE*, 6(12), e28716. doi:[10.1371/journal.pone.0028716](https://doi.org/10.1371/journal.pone.0028716)

Chapter 3

Decoding the direction of imagined visual motion using 7T ultra-high field fMRI

based on: Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A., & Goebel, R. (2016). Decoding the direction of imagined visual motion using 7T ultra-high field fMRI. *NeuroImage*, 125, 61–73.

and

Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A., & Goebel, R. (2016). Time-resolved searchlight analysis of imagined visual motion using 7T ultra-high field fMRI: Data on interindividual differences. *Data in Brief*, 7, 468–471.

We would like to thank Franziska Emmerling, Felix Dücker, Tom de Graaf, Lars Hausfeld, and Giancarlo Valente for their help.

ABSTRACT

There is a long-standing debate about the neurocognitive implementation of mental imagery. One form of mental imagery is the imagery of visual motion, which is of interest due to its naturalistic and dynamic character. However, so far only the mere occurrence rather than the specific content of motion imagery was shown to be detectable. In the current study, the application of multi-voxel pattern analysis to high-resolution functional data of 12 subjects acquired with ultra-high field 7 Tesla functional Magnetic Resonance Imaging allowed us to show that imagery of visual motion can indeed activate the earliest levels of the visual hierarchy, but the extent thereof varies highly between subjects. Our approach enabled classification not only of complex imagery, but also of its actual contents, in that the direction of imagined motion out of four options was successfully identified in two thirds of the subjects and with accuracies of up to 91.3% in individual subjects. A searchlight analysis confirmed the local origin of decodable information in striate and extra-striate cortex. These high-accuracy findings not only shed new light on a central question in vision science on the constituents of mental imagery, but also show for the first time that the specific sub-categorical content of visual motion imagery is reliably decodable from brain imaging data on a single-subject level.

3.1 INTRODUCTION

The nature and exact representation of complex mental imagery in the human brain is recognised as one of the keys to understanding creativity and intelligence (Ferguson, 1977; Kozhevnikov, Kozhevnikov, Yu, & Blazhenkova, 2013). If neuroimaging studies were able to reveal the neural correlates of mental imagery with a similar degree of accuracy as for perception, this may potentially allow for a better understanding of human creativity and intelligence since mental imagery seems to be a prerequisite for both (Miller, 1996).

One particularly naturalistic form of imagery is the imagination of visual motion. Visual motion is a crucial domain in daily life; the processing of visual motion is not only necessary to maintain a coherent percept of one's surrounding but also enables us to separate objects from each other and perceive depth. So far, research in cognitive neuroscience has mostly focused on static mental images such as orientations (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Harrison & Tong, 2009), shapes (Stokes, Saraiva, Rohenkohl, & Nobre, 2011; Stokes, Thompson, Cusack, & Duncan, 2009), objects (Cichy, Heinzle, & Haynes, 2012; Lee, Kravitz, & Baker, 2012; Reddy, Tsuchiya, & Serre, 2010), and scenes (M. R. Johnson & Johnson, 2014). Investigating motion imagery on the other hand goes beyond static and towards dynamic imagery processes and could, thus, enable a better understanding of sustained forms of imagery. Due to its naturalistic character, motion imagery is close to and directly translatable to actual action implementation, for instance in human machine interfacing. While neural activation patterns during visual short-term memory of perceived visual motion have been shown recently (Emrich, Riggall, LaRocque, & Postle, 2013), scientific knowledge about neural correlates of

motion imagery is still limited. This is due to the inherently private nature of imagery and inter-individual differences in personal strategies (Kozhevnikov, Hegarty, and Mayer, 2002; Kozhevnikov, Kosslyn, and Shephard, 2005; Poltrock and Brown, 1984; also note the recently described condition of congenital aphantasia: Zeman, Dewar, and Della Sala, 2015) and brain areas (Kosslyn, Ganis, & Thompson, 2001; Motes, Malach, & Kozhevnikov, 2008; Wraga, Thompson, Alpert, & Kosslyn, 2003) recruited during imagery. In addition, so far only the occurrence (categorical decoding), but not the specific content (sub-categorical decoding) of motion imagery has been decodable (Goebel, Khorram-Sefat, Muckli, Hacker, & Singer, 1998) and the question of whether the actual specific content of imagined motion is decodable, is still to be answered.

When it comes to the neural correlates of imagery, numerous findings show that perception and imagery share neural circuits (for a review see Kosslyn, Thompson, and Ganis, 2006). This suggests that also brain regions involved in visual motion imagery overlap with those involved in visual motion perception. The most important motion-responsive brain areas include direction-selective cells in V_1 (Hubel & Wiesel, 1968; McLean & Palmer, 1989; Movshon, Thompson, & Tolhurst, 1978), human area V_3A (Tootell et al., 1997; Van Oostende, Sunaert, Van Hecke, Marchal, & Orban, 1997), and the human motion complex (hMT+, Cheng, Fujita, Kanno, Miura, and Tanaka, 1995; Dupont, Orban, De Bruyn, Verbruggen, and Mortelmans, 1994; Tootell et al., 1995). Different directions of perceived visual motion can be decoded from several of these areas in the visual cortex (Kamitani & Tong, 2005). More recently, the axis-of-motion selective columnar structure in the middle temporal area (MT) that is known from animal studies (Albright, Desimone, & Gross, 1984; Albright, 1984; Diogo, Soares, Koulakov, Albright, & Gattass, 2003; Dubner & Zeki, 1971) was

shown in humans using 7 Tesla (T) ultra-high field functional magnetic resonance imaging (fMRI; Zimmermann et al., 2011). In light of these previous studies, V₁, V_{3A}, and hMT+ are the most likely candidates for the decoding of visual motion imagery.

In order to optimally decode information regarding imagery within these regions, high-field imaging and multivariate analyses seem promising. The recent developments of 7 T ultra-high field fMRI for humans has brought improvements in spatial specificity (Uğurbil et al., 2003; Uludağ, Müller-Bierl, & Uğurbil, 2009) and increases in signal-to-noise ratio (SNR; Vaughan et al., 2001) over conventional 3 T fMRI. Additionally, with higher magnetic field strengths the influence of large draining veins on the blood-oxygen level dependent (BOLD) signal becomes smaller, while the signals from small vessels and capillaries become better detectable (Shmuel, Yacoub, Chaimow, Logothetis, & Uğurbil, 2007). This further enhances the functional specificity. Seminal studies have leveraged these benefits to explore the functional organisation of the visual system at the level of cortical columns (Cheng, Waggoner, & Tanaka, 2001; Goodyear & Menon, 2001; Menon, Ogawa, Strupp, & Uğurbil, 1997; Yacoub, Harel, & Uğurbil, 2008; Yacoub, Shmuel, Logothetis, & Uğurbil, 2007). These structures are well known from animal studies (Bonhoeffer & Grinvald, 1991; Hubel & Wiesel, 1968; Mountcastle, 1957; Tanaka, 1996) and seem to be a crucial functional unit providing a key to understand computational mechanisms in early sensory areas and beyond. In addition to hardware advancements, new data analysis techniques like multi-voxel pattern analysis (MVPA; for an introduction see Pereira, Mitchell, and Botvinick, 2009) and information-based functional brain mapping (Kriegeskorte, Goebel, & Bandettini, 2006) enable the extraction of more spatially distributed information from fMRI data. However, there is an ongoing debate on the spatial scale of

the signals that MVPA picks up. There are controversial findings on whether large-scale biases can explain decoding performances of MVPA decoding studies (Op de Beeck, 2010a) rather than signals from a fine-grained columnar organisation of cortex (Swisher et al., 2010).

Together, ultra-high field fMRI and new analysis algorithms allow one to overcome the resolution boundaries given by conventional neuroimaging methods. Here we apply these methodological advances to the study of visual motion imagery.

The current study aims to decode the specific content of visual motion imagery, accounting for inter-individual differences in imaging strategies and neural activation patterns. In this way, we combine the described advantages of high-field brain imaging and multivariate analyses to decode not simply the occurrence (categorical decoding), but the specific content (sub-categorical decoding) of self-generated motion imagery. Mental imagery during the experiment was well controlled (participants were trained over multiple sessions with the imagery task) and as neutral as possible (the type of visual imagery was chosen to be abstract enough to not interfere with personal memories etc.), while still being completely self-generated (i. e., the experimental paradigm did not include any visual stimulation that could elicit bottom-up activations of the visual system or reactivation following, e. g., a motion stimulus of a previous trial). Therewith we aimed to trigger low-level neural correlates of motion imagery and, thus, expected decodable information on the imagined motion direction in early visual areas (Kamitani & Tong, 2006) and MT (Goebel et al., 1998). Furthermore, in an exploratory way we attempted to link inter-individual differences in the brain activation patterns recruited during motion imagery to individual cognitive imagery styles assessed with self-report questionnaires.

3.2 METHODS

SUBJECTS 15 healthy fMRI-experienced subjects (six females; age: 27.4 ± 6.3 years) with normal or corrected-to-normal vision volunteered in this study. They gave written informed consent and were paid for their participation. All procedures were conducted with approval from the local Ethical Committee of the Faculty of Psychology and Neuroscience at Maastricht University. Two subjects had to be excluded from the analyses because no significant activations could be detected based on a localizer scan (see below). One subject was excluded due to discomfort during the scanning resulting in problems to follow the experimental instructions. All subjects were students or employees at Maastricht University.

PROCEDURE In the experiment, subjects had to imagine dots that moved in one of four directions. Subjects were pseudo-randomly assigned to one of two groups. The first group imagined motion directions left, right, up, and down whereas the second group imagined the four diagonal motion directions. Subjects heard a cue on which specific motion direction they should imagine during the upcoming trial. These cues were auditorily presented numbers (1–4) and were associated to the four motion directions in a counterclockwise (subjects 1–9) or clockwise order (subjects 10–15). The orientation of this number-motion direction association (e. g., “1” = right versus “1” = up) was systematically varied across subjects.

TRAINING SESSIONS All subjects attended three training sessions during the week before the scanning session. In each training session subjects followed an adaptive visual imagery training (one run in the first session, two runs in the remaining sessions; see

below) and completed two runs of the experimental task (see below) with shortened resting periods. Each training session lasted between 30 and 45 minutes.

ADAPTIVE VISUAL IMAGERY TRAINING Coherently moving white dots were presented on the screen together with an auditory cue and faded out after 4000 ms leaving only a white fixation dot. When indicated by a change of fixation colour, subjects were instructed to imagine the faded stimulus as vividly as possible. They were asked to press a button as soon as they had a clear picture or 'movie' of the stimulus in their mind and to hold the button down for as long as they 'saw' this imagined stimulus. If subjects released the button before 8000 ms had passed the stimulus faded in again and they could try again by pressing the button, which faded the stimulus out. After successfully imagining the stimulus for more than 8000 ms the trial ended and was followed by a 5-point Likert scale (1: "No image at all, you only 'know' that you are thinking of the object" to 5: "Perfectly clear and vivid as normal vision") to indicate the clearness of the imagery in the preceding trial. After the adaptive visual imagery training the subjects reported the used imagery strategies to the experimenter.

EXPERIMENTAL TASK In the beginning and in the middle of each run, moving dots were presented as in the training task together with a white fixation dot. This was done in order to facilitate the subjects' recall of the precise stimulus configuration and the number-motion direction association. The four motion directions that had to be imagined by each group were presented for 4000 ms each together with the associated auditory cue (read-out number; from "1" to "4"). After this reference phase and a resting period of 12 s the first block of four trials started (see [Figure 3.1](#)).

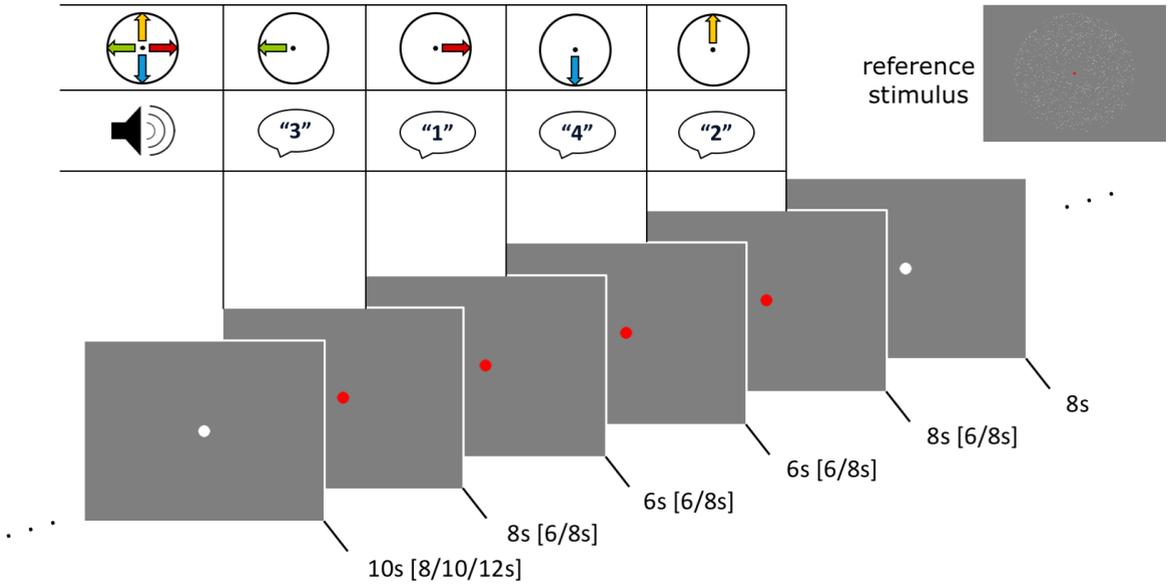


Figure 3.1 Experimental task. Visual stimulation is shown exemplarily for one block of four trials (red fixation dots) with preceding and succeeding resting blocks (white fixation dots). Durations are shown below each screenshot (possible jittered duration are given in brackets). In the table at the top the motion directions with their corresponding audio cues (that were played to the subject at the start of a trial) are shown (number-motion direction association varied between subjects). At the top right one frame of the reference stimulus is shown representing the reference phase.

We chose to present trials in blocks to maximise the number of trials while allowing subjects to recover from the cognitively demanding task of motion imagery during resting periods. When the fixation dot turned red and the first auditory cue was presented the subjects had to start to imagine the associated motion direction as vividly as possible. The following trials were indicated by the presentation of auditory cues and the subjects had to start to imagine the associated motion direction immediately. Trials lasted for six or eight seconds. The end of the last trial in a block (and thus the cue to stop all imagery) was indicated by the fixation dot turning white again. After a resting period of eight seconds, a 5-point Likert scale was presented. Subjects indicated the clearness of the imagery in the preceding trial block by moving an arrow via button press. After another resting period with a jittered duration of eight, ten, or twelve seconds the next block started. Subjects were instructed to fixate the fixation dot throughout the whole experiment. Each run consisted of ten blocks and the subjects completed four runs in the scanner (total duration of experimental runs approximately 40 minutes; 40 trials per condition).

STIMULI Visual stimulation was created with PsychoPy (version 1.78.01; Peirce, 2007) and, in the scanning session, projected on a frosted screen at the top end of the scanner bed using an LCD projector (Panasonic, No PT-EZ57OEL; Newark, NJ, USA). Responses were collected through an MR compatible button box (Current Designs, 8-button response device, HHSC-2x4-C; Philadelphia, USA).

VIVIDNESS OF VISUAL IMAGERY QUESTIONNAIRE After recruitment (before the training sessions) subjects filled in the Vividness of Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). This self-report questionnaire measures subjective clearness

and vividness of imagined objects and scenes with 16 items. The vividness ratings for each imagined item are given on a scale from 1 (“No image at all, you only ‘know’ that you are thinking of the object”) to 5 (“Perfectly clear and vivid as normal vision”). After scanning subjects filled in the VVIQ for a second time.

OBJECT-SPATIAL IMAGERY AND VERBAL QUESTIONNAIRE

After the scanning session, subjects were contacted to fill out the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ; Blazhenkova and Kozhevnikov, 2009). The OSIVQ is a self-report questionnaire consisting of three scales for “object”, “spatial”, and “verbal” cognitive styles during mental imagery measured by 15 items each. In each item a statement is rated on a scale from 1 (“totally disagree”) to 5 (“totally agree”). We calculated the score on each scale for every subject as described in the original paper (Blazhenkova & Kozhevnikov, 2009).

EYE-TRACKING To check for eye-movements related to the different directions of imagined motion we recorded eye movements during the scanning session for four subjects (9–12) using an MR-compatible eye-tracker (Real Eye Nano; Avotec, Inc.; Stuart, FL, USA). Eye-tracking data were analysed using custom code in MATLAB (version 2013a; The Mathworks Inc., Natick, MA, USA), code from the “EYE-EEG extension” toolbox (Dimigen, Sommer, Hohlfeld, Jacobs, and Kliegl, 2011; <http://www2.hu-berlin.de/eyetracking-eeeg>) to detect saccades based on the algorithm by Engbert and Mergenthaler (2006), and code from the CircStat toolbox (Berens, 2009). To assess the statistical similarity of saccade directions across trials with different imagined directions we used a non-parametric multi-sample test for equal median directions as implemented in the CircStat toolbox under the null hypothesis

that saccade directions did not differ between different directions of imagined motion. We also computed the circular-circular correlation between the direction of imagined motion and the saccade direction and tested it for significance. Furthermore, we analysed the raw eye gaze position data during trials. After discarding data points during eye blinks (pupil aspect ratio lies outside a confidence interval of ± 2 SD around the average pupil aspect ratio in each trial) we fitted a Minimum Volume Enclosing Ellipse (Moshtagh, 2005) to the XY data of each trial. Then we statistically tested the similarity of the ellipse rotations across trials and the circular-circular correlation with the different imagined directions employing the same methods we used for the saccade data (see above).

SCANNING SESSION First, each subject completed a short 4-block practice version of the experimental task outside the scanner. At the beginning of the scanning session, we recorded an hMT+ localizer scan as in Zimmermann et al. (2011) followed by a population receptive field (pRF) retinotopy mapping scan (Dumoulin & Wandell, 2008).

MRI ACQUISITION Images were acquired with a Siemens Magnetom 7 T scanner (Siemens; Erlangen, Germany) and a 32-channel head-coil (Nova Medical Inc.; Wilmington, MA, USA). An anatomical dataset was acquired with a T₁-weighted magnetisation prepared rapid acquisition gradient echo (3D-MPRAGE) sequence (256 sagittal slices, matrix = 384×384 , voxel size = $0.6 \times 0.6 \times 0.6$ mm³). To correct for intensity inhomogeneities an additional gradient echo proton-density (GE-PD) dataset (same parameters as 3D-MPRAGE) was acquired subsequently.

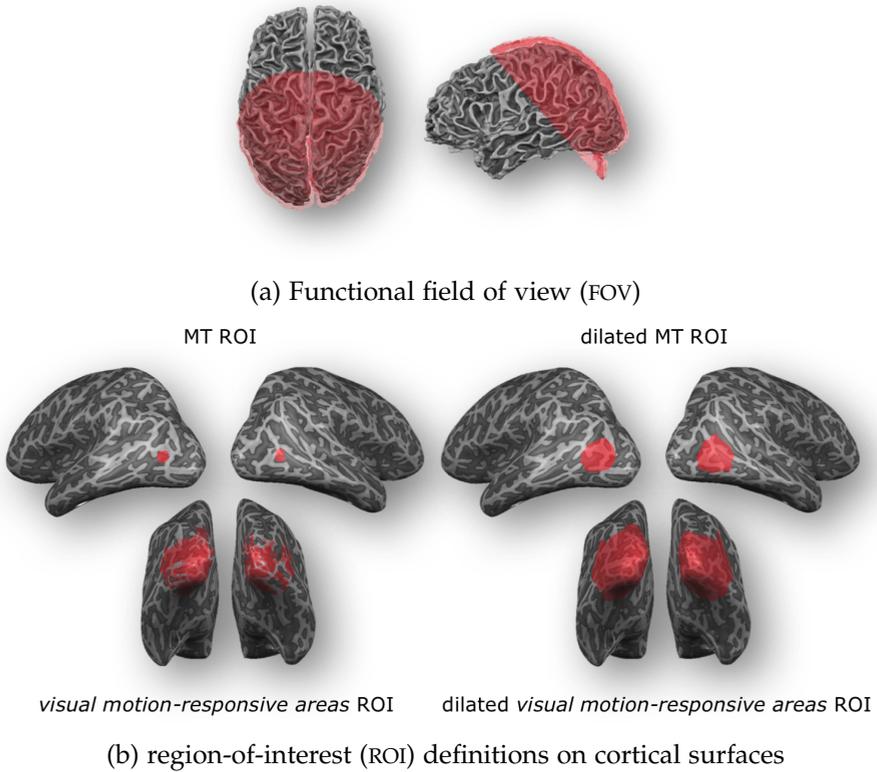


Figure 3.2 Functional field of view (FOV) and region-of-interest (ROI) definitions on cortical surfaces. **(a)** The FOV is shown exemplarily in subject 2 for the functional images (red) from two perspectives (left: top view; right: left-hemispheric view). **(b)** The MT ROI and the *visual motion-responsive areas* ROI are exemplarily shown in the normal (left) and dilated (right) versions in subject 2.

High-resolution functional images were obtained using gradient echo (T_2^* weighted) echo-planar imaging (EPI) with the following parameters: echo time (TE) = 23 ms, repetition time (TR) = 2000 ms, generalised auto-calibrating partially parallel acquisitions (GRAPPA) factor = 2, multi-band factor = 2, nominal flip angle = 70° , number of slices = 54, matrix = 130×130 , voxel size = $1.1 \times 1.1 \times 1.1 \text{ mm}^3$. The field-of-view included occipital, temporal, and parietal areas but did not cover large parts of frontal cortex (see [Figure 3.2a](#)). To correct for EPI distortions additional functional volumes (five volumes in the encoding direction and five volumes with a reversed encoding direction) were acquired right after the GE-PD dataset.

IMAGING DATA PREPROCESSING Functional and anatomical images were analysed using BrainVoyager QX (version 2.8; Brain Innovation; Maastricht, The Netherlands), custom code in MATLAB (version 2013a; The Mathworks Inc.; Natick, MA, USA), and PyMVPA (version 2.3; Hanke et al., 2009). Anatomical images were corrected for bias field inhomogeneities by dividing the 3D-MPRAGE images by the GE-PD images and interpolated to a nominal voxel size of 0.55 mm isotropic to match a multiple of 2 of the resolution of the functional data. The detection of the white/grey matter boundary was conducted with the largely automatic segmentation tools of BrainVoyager QX. These tools perform a region-growing method that analyses intensity histograms and subsequently correct topological errors in the detected borders to finally reconstruct the cortical surfaces (Goebel, Esposito, & Formisano, 2006; Kriegeskorte & Goebel, 2001).

In order to correct for distortions in the echo-planar images we recorded five functional volumes of normal and reversed phase encoding. In these pairs of images distortions go in opposite direc-

tions and we used them to estimate the susceptibility-induced off-resonance field using a method similar to that described in Andersson, Skare, and Ashburner (2003) as implemented in FSL (Smith et al., 2004). After performing 3D rigid body motion correction of the remaining functional runs (aligning all subsequent runs to the first functional run) the estimated off-resonance field was used to correct for EPI distortions. Furthermore, functional data were high-pass filtered using a general linear model (GLM) Fourier basis set of two cycles sine/cosine per run (including linear trend removal). Functional runs were co-registered to the individual anatomical scan with an affine (9 parameter) transformation.

REGION-OF-INTEREST DEFINITION Regions of interest (ROIs) were then defined using data from the hMT+ localizer scan and the pRF retinotopy projected onto an inflated surface reconstruction. Area MT was distinguished from medial superior-temporal area (MST) in the hMT+ complex (Huk, Dougherty, & Heeger, 2002; Zimmermann et al., 2011) by thresholding the respective contrast at $p < .001$ and restricting the resulting activation patch to 150 vertices per hemisphere (for creating the MT ROI). A second ROI was created by applying the same contrasts but thresholding at $p < .05$ and including all active patches in visual cortices (for creating the “visual motion-responsive areas” ROI) in each hemisphere. For extended versions of these ROIs the patches were dilated by a static factor (adding neighbour vertices at the patches’ boundaries; 20-fold dilation for the MT ROI, 10-fold dilation for *visual motion-responsive areas* ROI) to include possibly contributing neighbouring regions (see [Figure 3.2b](#) and [Table 3.2](#)). Areas V₁ through visual area 4 (V₄) were delineated in the polar angle map of the surface-based pRF analysis in both hemispheres. All surface patches were transformed back into volume space (from -1 mm till +3 mm from

the grey/white matter segmentation boundary) to create the final volume ROIs.

MULTI VOXEL PATTERN ANALYSES Each experimental run was z-scored to eliminate signal offsets and variance differences between runs. After masking the experimental data with the individually defined ROIs the data were split into training and testing datasets. We employed a leave-one-run-out splitting procedure to be able to cross-validate the classification performance. For each split the 1000 voxels (within the respective ROI) with the highest F-values in the respective training data were selected as features. The F-values were computed as the standard fraction of between and within class variances (omnibus test). This was done to reduce the high number of voxels within ROIs in 7T fMRI scans and in order to keep the number of features constant between subjects and ROIs. For each voxel and each trial, we extracted the average of six seconds (3 TRs) as features. Averages were computed in a time window from four to ten seconds after trial onset. The extracted features were then used for a one-vs-one 4-class classification (predicted classes were chosen based on the maximum number of votes in all binary classifications) using a linear support vector machine (SVM; LIBSVM implementation in PyMVPA; Chang and Lin, 2011). We repeated the whole analysis 1000 times with scrambled labels to obtain a distribution under the null hypothesis and tested the probability of the real classification accuracies against this distribution. To assess a group level statistic, we tested the real classification performances of each subject against the individual average permutation classification accuracy by means of a Wilcoxon Signed-Ranks test. We chose a non-parametric test as a normal distribution of the decoding accuracies cannot be assumed. Furthermore, we computed Spearman rank-order correla-

tions between mean classification accuracies (across all ROIs) and the OSIVQ scores (*object*, *spatial*, and *verbal* scales). We tested these correlations for significance after an FDR-correction for multiple comparisons to reveal possible influences of cognitive imagery styles on the classification performance. To assess any differences between the group of subjects imagining horizontal/vertical directions and the group of subjects imagining diagonal directions we tested mean classification accuracies (across all ROIs) by means of a Mann-Whitney U test across the two groups.

SEARCHLIGHT ANALYSIS To assess the spatial distribution of brain areas involved in the mental imagery task without restriction to defined ROIs we performed a searchlight analysis (Kriegeskorte et al., 2006). A sphere with a radius of four voxels was moved through the cortical ribbon (so that the spheres central voxel always lay within -1 mm to +3 mm from the grey/white matter segmentation border) and defined a feature set of 257 voxels (four voxel radius). For each voxel within each sphere and each trial we extracted the average of 6 s (3 TRs) as features. Averages were computed in six different but overlapping time windows starting at -2, 0, 2, 4, 6, and 8 s (i. e., -1, 0, 1, 2, 3, and 4 TRs) with respect to the trial onset. These features that was in turn analysed with the MVPA pipeline outlined above (without voxel pre-selection and permutation testing). The resulting classification accuracies were tested for significance by means of FDR-corrected Chi-square tests of the confusion matrix and projected onto the inflated surface reconstruction. In the results section, searchlight results are shown for the time window (four until ten seconds after trials onset). Volumetric significance maps for all time windows were MNI space transformed (see below) and uploaded to the [NeuroVault.org](https://neurovault.org) repository.

Furthermore, we mapped the direction selectivity of these brain areas. A preference map was computed by comparing the four t-value maps of single-direction contrasts and subsequently masked by the significant ($p < .05$) areas in the Chi-square searchlight map.

MNI SPACE TRANSFORMATION In order to make individual searchlight accuracy maps available for inspection, individual anatomies were transformed to MNI-ICBM 152 space (Mazziotta, Toga, Evans, Fox, Lancaster, Zilles, Woods, Paus, Simpson, Pike, Holmes, Collins, Thompson, MacDonald, Iacoboni, Schormann, Amunts, Palomero-Gallagher, Geyer, Parsons, Narr, Kabani, Le Goualher, et al., 2001) using BrainVoyager 20.0 and the ICBM 452 template (Mazziotta, Toga, Evans, Fox, Lancaster, Zilles, Woods, Paus, Simpson, Pike, Holmes, Collins, Thompson, MacDonald, Iacoboni, Schormann, Amunts, Palomero-Gallagher, Geyer, Parsons, Narr, Kabani, Goualher, et al., 2001). The individual significance maps for the six different time windows were resampled to a resolution of $1 \times 1 \times 1 \text{ mm}^3$ voxel size and transformed into MNI space using the individual transformation matrices obtained in the previous step. The resulting maps were named according to the subject ("S01" through "S12") and the time window ("_shift_-1" through "shift_4") and uploaded to [NeuroVault.org](http://neurovault.org/collections/961/) (<http://neurovault.org/collections/961/>).

UNIVARIATE ANALYSIS We performed a GLM analysis for two single subjects that showed very high MVPA decoding accuracies as significant univariate results are expected in such cases. Furthermore, significant univariate results and corresponding maps provide evidence that the MVPA classifier did not pick up on some artifactual or confounding signal rather than brain activity related to the motion imagery task. We used linear predictors for each exper-

imental condition convolved with a standard two-gamma hemodynamic response function. We then computed contrasts for each of the directions of imagined motion against all other directions. We plotted four corresponding t-maps that were thresholded at $q < .05$ (FDR corrected) on the inflated cortical surfaces.

3.3 RESULTS

BEHAVIOURAL DATA VVIQ scores for all subjects are shown in [Figure 3.3](#). Scores did not change significantly between before the training and after the scanning session (paired t-test; $t(11) = 0.12$; $p = .903$). OSIVQ scores on the three scales for all subjects are shown in [Figure 3.4](#). The strategies initially reported by subjects are shown in [Table 3.1](#). In the third training session most subjects reported to not use particular strategies anymore but visualise the reference stimulus directly.

The training sessions improved the subjective ratings for the vividness of the imagery in most subjects (see [Figure 3.5](#)). On group level, a paired t-test between behavioural data from the first and the last training session revealed a significant improvement for the ratings in the adaptive visual imagery training ($t(11) = -3.64$; $p = .004$) and the experimental task ($t(11) = -3.45$; $p = .005$). Data from the ratings given during the scanning sessions were unfortunately lost due to a software malfunction and, hence, could not be used for data analysis.

MULTI VOXEL PATTERN ANALYSES Volumes of all defined ROIs are shown in [Table 3.2](#). Group-level statistics revealed significant classification accuracies for all defined ROIs (MT ($W = 9$; $p = .019$), dilated MT ($W = 0$; $p = .002$), *visual motion-responsive areas* ($W = 0$;

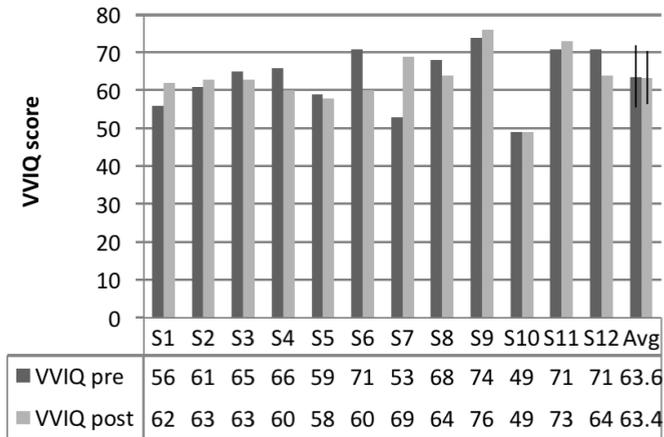


Figure 3.3 Vividness of Visual Imagery Questionnaire (vviq) scores. vviq scores from before the training (pre) and after the scanning session (post) are shown for all subjects (S1 through S12). The last column shows the group averages (Avg); error bars indicate one standard deviation.

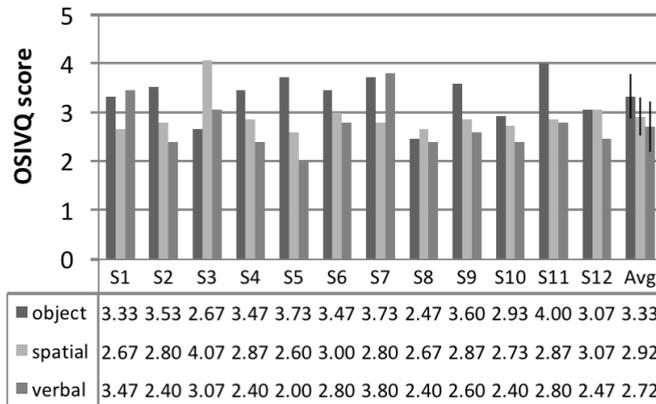


Figure 3.4 Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) scores. OSIVQ scores for the three different scales “object”, “spatial”, and “verbal” are shown for all subjects (S1 through S12). The last column shows the group averages (Avg); error bars indicate one standard deviation.

SUBJECT	STRATEGY
1	upwards: bubbles rising in a glass; downwards: snow falling (blizzard); other two directions: city lights viewed from an airplane
2	downwards: rain; no other specific strategy
3	up/down: tetris game; left/right: tennis match
4	rain/bubbles while visualizing the aperture to have a guidance for the imagery
5	swimming tadpoles
6	no particular strategy
7	no particular strategy
8	focus on the dot pattern; no other particular strategy
9	upwards: bubbles rising in a glass; downwards: snow falling; left/right: buffalo herd running
10	try to keep the almost faded out dots in mind
11	try to keep the almost faded out dots in mind
12	upwards: water hose; downward: rain

Table 3.1 Strategies used by subjects as reported during the first training session.

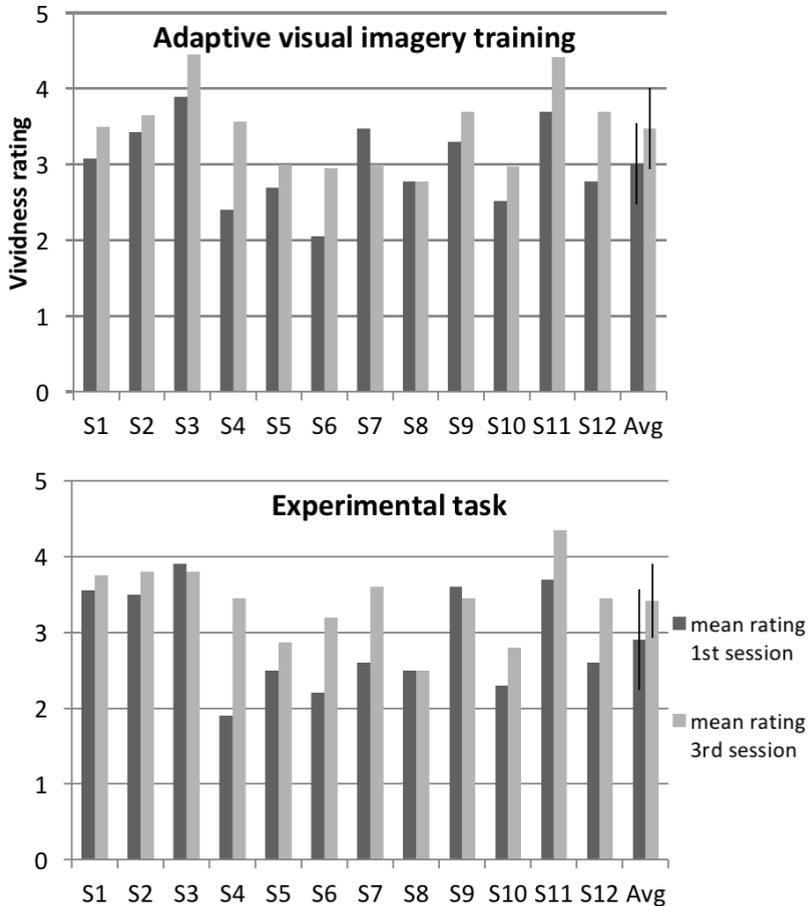


Figure 3.5 Mean vividness ratings for adaptive visual imagery and experimental task in training sessions. Mean ratings for the vividness of imagery (1: “No image at all, you only ‘know’ that you are thinking of the object” to 5: “Perfectly clear and vivid as normal vision”) are shown for all subjects (S1 through S8) for the first and the last training session in the adaptive visual imagery training (left) and the experimental task (right). The last columns show the group averages (Avg); error bars indicate one standard deviation.

$p = .002$), dilated *visual motion-responsive areas* ($W = 0$; $p = .002$), V_1 ($W = 12$; $p = .034$), secondary visual area (V_2) ($W = 5$; $p = .008$), tertiary visual area (V_3) ($W = 6$; $p = .01$), and V_4 ($W = 3$; $p = .005$) ROIs. The single-subject level MVPAs revealed significant classification accuracies for two subjects in the MT ROI, for two subjects in the dilated MT ROI, for five subjects in the V_1 , V_2 , V_3 , and V_4 ROIs, for six subjects in the dilated *visual motion-responsive areas*, and for eight subjects in the *visual motion-responsive areas* ROIs, respectively (see [Figure 3.6](#)). In subject 7 classification accuracies of up to 91.25% (V_3 ROI; chance level at 25%) were observed. Mean classification accuracies (across all ROIs) did not differ significantly between the group of subjects imagining horizontal/vertical motion directions and the group of subjects imagining diagonal motion directions ($U = 10$; $p = .128$).

BEHAVIOURAL CORRELATIONS No correlation between OSIVQ scores and classification accuracies was significant ($r_{\text{Object}} = .365$ ($p = .243$); $r_{\text{Spatial}} = -.304$ ($p = .337$); $r_{\text{Verbal}} = -.185$ ($p = .564$); see [Figure 3.7](#)).

SEARCHLIGHT ANALYSIS Maps showing cortical areas with significant searchlight accuracies showed inter-individual differences (see [Figure 3.8](#)). In subject 7 large parts of the early visual system showed significant decoding accuracies (see [Figure 3.9a](#)). In subject 5 mainly V_2 , V_3 , and V_4 areas going into the foveal confluence showed significant decoding accuracies (see [Figure 3.9b](#)). In subject 6 areas V_2L , V_3L , and V_4L (and beyond) showed significant decoding accuracies in the left hemisphere. In the right hemisphere area V_4L showed significant decoding accuracies (see [Figure 3.9c](#)). In subject 8 a superior parietal region showed significant decoding accuracies.

SUBJECT	MT	MT D	MA	MA D	V1	V2	V3	V4
1	1026 (1.366)	8496 (11.308)	20958 (27.895)	40373 (53.736)	3343 (4.45)	4087 (5.44)	3173 (4.223)	2900 (3.86)
2	914 (1.217)	7497 (9.979)	32198 (42.856)	53566 (71.296)	4489 (5.975)	4541 (6.044)	4216 (5.611)	3621 (4.82)
3	1161 (1.545)	8404 (11.186)	28763 (38.284)	58326 (77.632)	5014 (6.674)	4315 (5.743)	4388 (5.84)	3334 (4.438)
4	1290 (1.717)	9344 (12.437)	27410 (36.483)	53552 (71.278)	2807 (3.736)	3708 (4.935)	3370 (4.485)	3076 (4.094)
5	1023 (1.362)	7584 (10.094)	41955 (55.842)	71583 (95.277)	4547 (6.052)	5049 (6.72)	4278 (5.694)	3678 (4.895)
6	986 (1.312)	8105 (10.788)	35565 (47.337)	63736 (84.833)	3808 (5.068)	3746 (4.986)	4088 (5.441)	3713 (4.942)
7	998 (1.328)	8145 (10.841)	38272 (50.94)	72269 (96.19)	5210 (6.935)	6244 (8.311)	6497 (8.648)	4535 (6.036)
8	1160 (1.544)	8486 (11.295)	23668 (31.502)	53451 (71.143)	3936 (5.239)	5032 (6.698)	4220 (5.617)	2912 (3.876)
9	990 (1.318)	8473 (11.278)	43215 (57.519)	75581 (100.598)	3596 (4.786)	3135 (4.173)	3034 (4.038)	3478 (4.629)
10	1013 (1.348)	8458 (11.258)	20816 (27.706)	42934 (57.145)	3820 (5.084)	3492 (4.648)	3016 (4.014)	2567 (3.417)
11	1003 (1.335)	7670 (10.209)	22128 (29.452)	44265 (58.917)	4226 (5.625)	3910 (5.204)	3431 (4.567)	3417 (4.548)
12	1012 (1.347)	8367 (11.136)	21256 (28.292)	52395 (69.738)	4500 (5.99)	3364 (4.477)	3163 (4.21)	3347 (4.455)
Avg	1048.0 (1.395)	8252.417 (10.984)	29683.67 (39.509)	56835.92 (75.649)	4108.0 (5.468)	4218.583 (5.615)	3906.167 (5.199)	3381.5 (4.501)
Std	98.825 (0.132)	484.973 (0.645)	8028.314 (10.686)	11298.38 (15.038)	663.526 (0.883)	846.353 (1.126)	933.923 (1.243)	483.801 (0.644)

Table 3.2 Volumes of defined Regions of interest (ROIs). Numbers of voxels (voxel size = $1.1 \times 1.1 \times 1.1 \text{ mm}^3$) for all defined ROIs are shown for all subjects. In parentheses volumes are shown converted to (rounded) cm^3 . MT d: 20-fold dilated MT ROI; ma: *visual motion-responsive areas* ROI; ma d: 10-fold dilated *visual motion-responsive areas* ROI. Avg: Average (across subjects); Std: Standard deviation (across subjects).

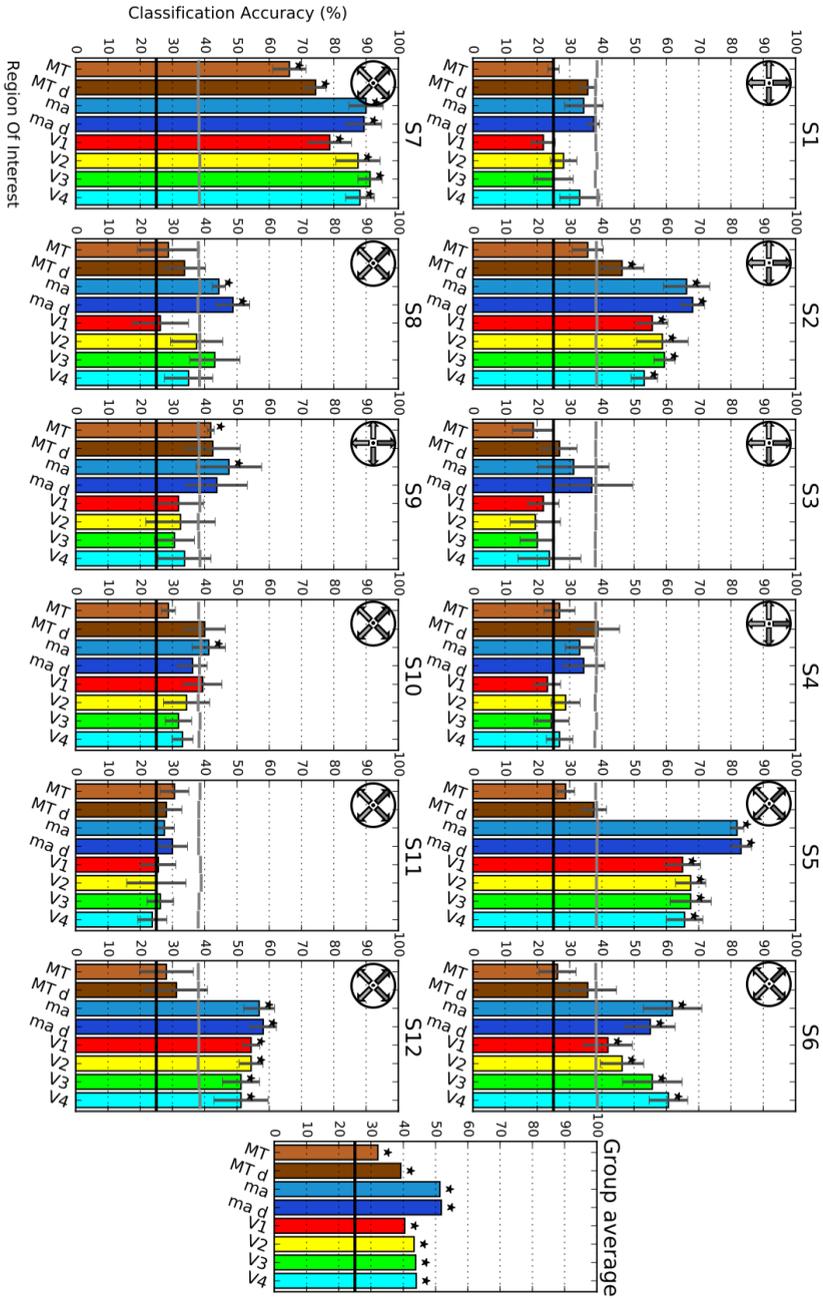


Figure 3.6 Classification accuracies for imagery of visual motion in four different directions. Average classification accuracies across cross-validations for all subjects (S1-S12) are shown in eight different ROIs. Error bars show one standard deviation of the accuracies' distribution across cross-validations. The black horizontal lines represent the chance level (25%); the grey horizontal lines show the 95th percentile of each 1000 permutations test. Asterisks in single subjects indicate significant accuracies ($p < .05$) as assessed by testing against 1000 permutation. Asterisks in the group average indicate significant accuracies as assessed by Wilcoxon Signed-Rank tests against the average classification accuracy of 1000 permutations. Orientation wheels at the right show the directions of motion that were imagined by the respective subjects in each row. MT: 20-fold dilated MT ROI; ma: *visual motion-responsive areas* ROI; ma d: 10-fold dilated *visual motion-responsive areas* ROI.

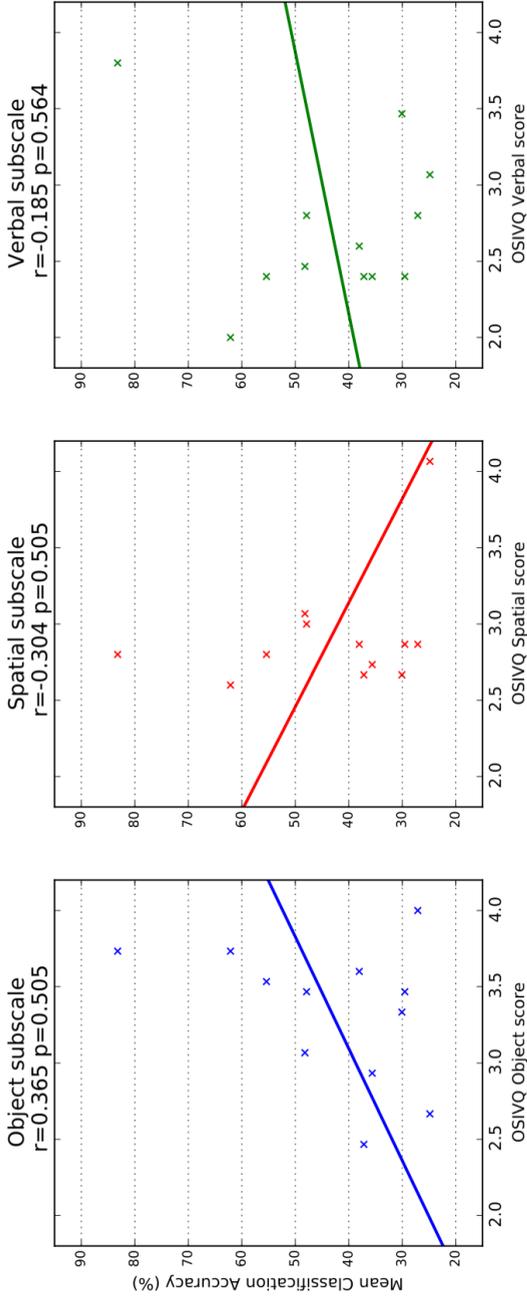


Figure 3.7 Scatter plots for Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) scores and average classification accuracies (across all ROIs). Scores from the three OSIVQ scales “object” (blue), “spatial” (red), and “verbal” (green) are plotted against the average classification accuracies for each subject. The coloured lines show linear regressions for each OSIVQ scale. Correlations and corresponding p-values are shown in titles of each subplot.

Preference maps revealed individually distinct patterns. In subjects 5 and 7 a difference in preference from central to peripheral parts of the visual field (following the eccentricity tuning) was observed.

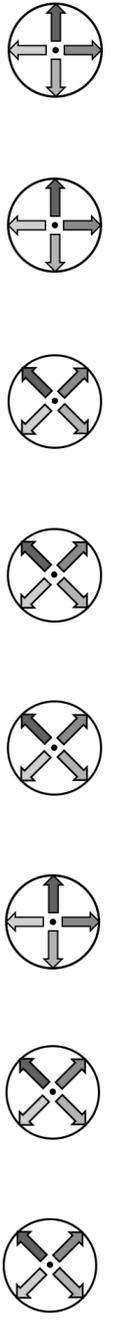
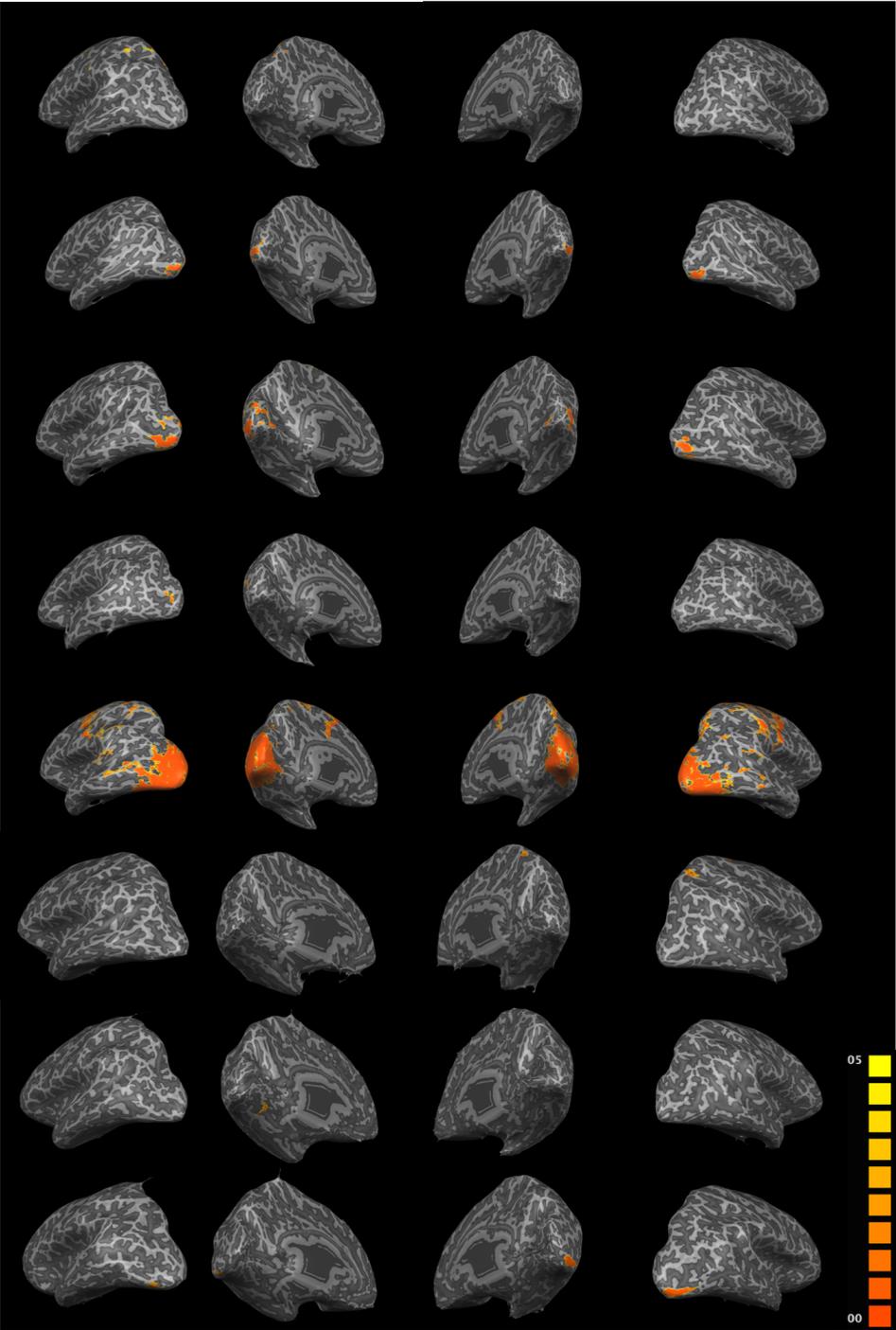
Volumetric maps showing p-values for FDR-corrected Chi-square tests of the confusion matrices of each searchlight sphere for all subjects and time windows were uploaded to [NeuroVault.org](http://neurovault.org/collections/961) (<http://neurovault.org/collections/961>).

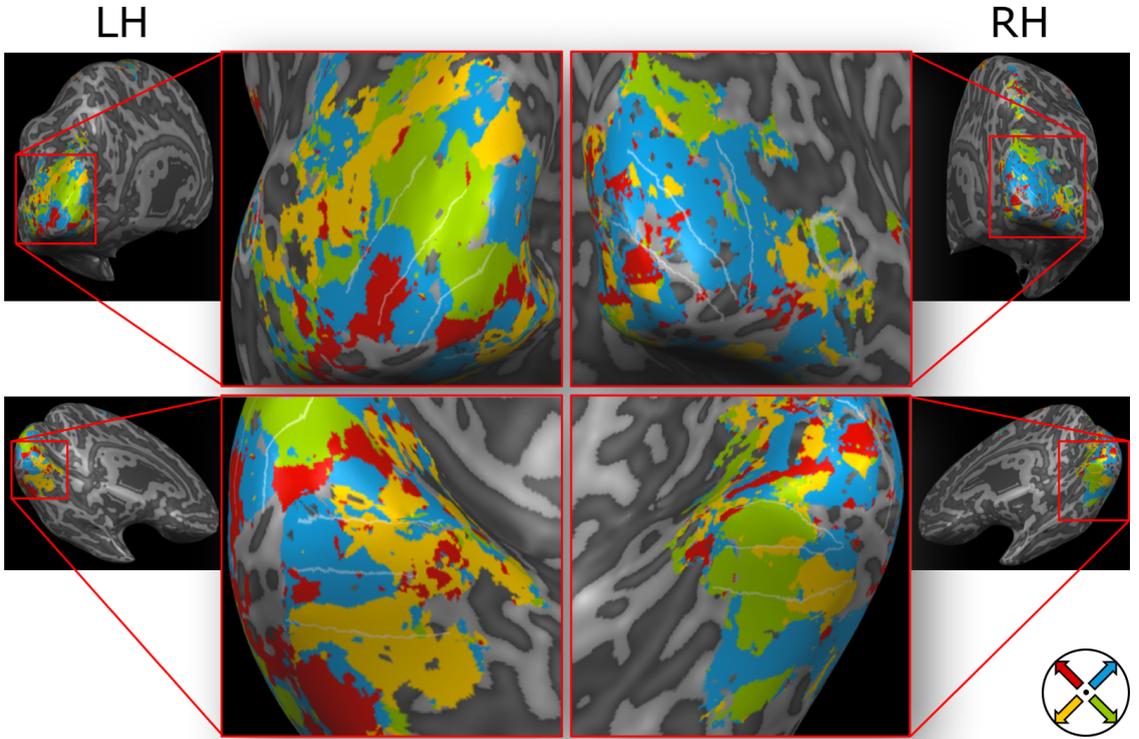
Figure 3.8 Searchlight maps for eight subjects (page 91). Significant areas in the searchlight map ($p < .05$ FDR corrected; cluster-thresholded at 50 mm^2) in the two inflated hemispheres of subjects 1, 2, 5, 6, 7, 9, 10, and 12 (there were no significant patches on the remaining subjects). Warmer colours (orange > yellow) indicate lower p-value. On the right the motion directions that had to be imagined are shown for each subject. LH: left hemisphere; RH: right hemisphere; lat.: lateral view; med.: medial view.

LH lat. LH med. RH med. RH lat.

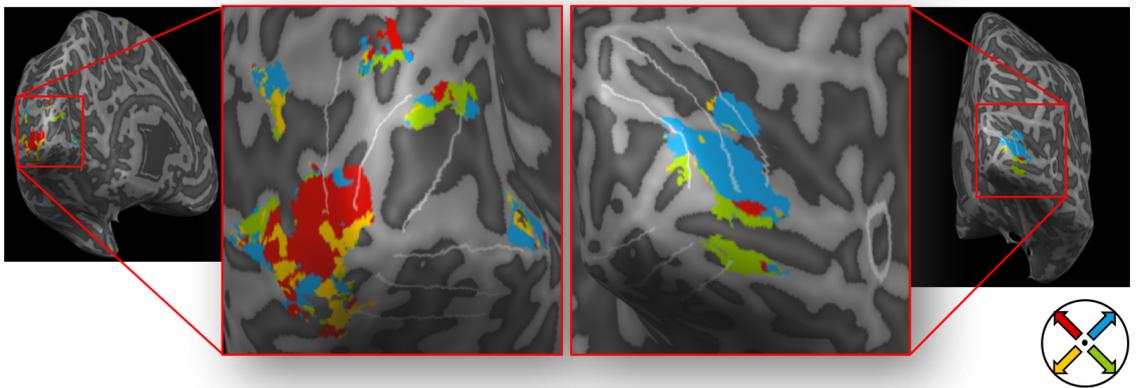
Subject

1
2
5
6
7
9
10
12

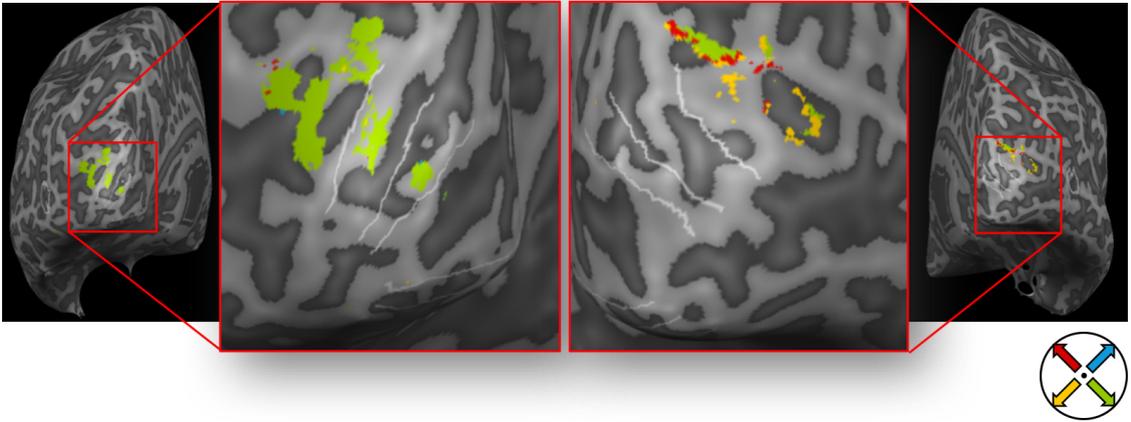




(a) Searchlight-thresholded univariate preference maps for subject 7



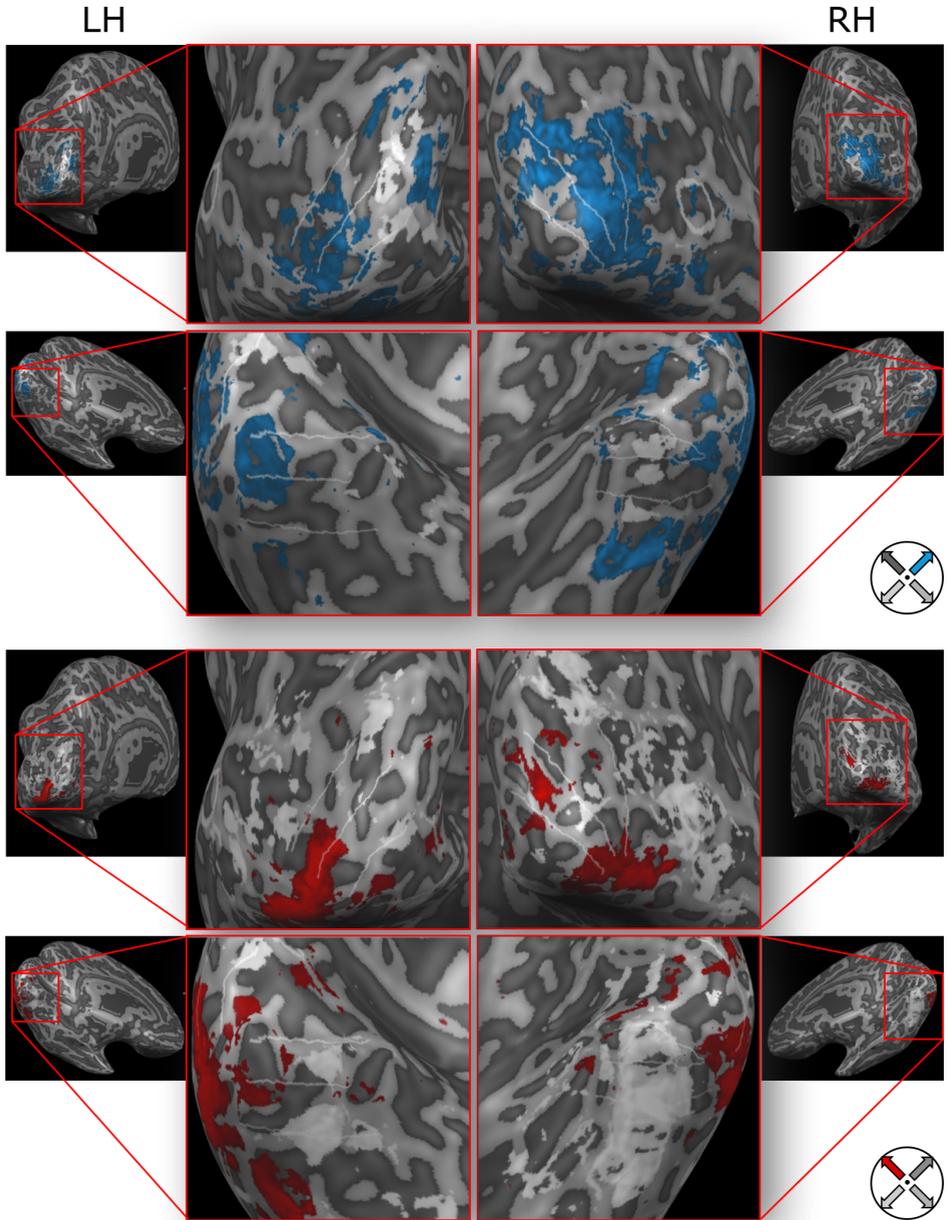
(b) Searchlight-thresholded univariate preference maps for subject 5

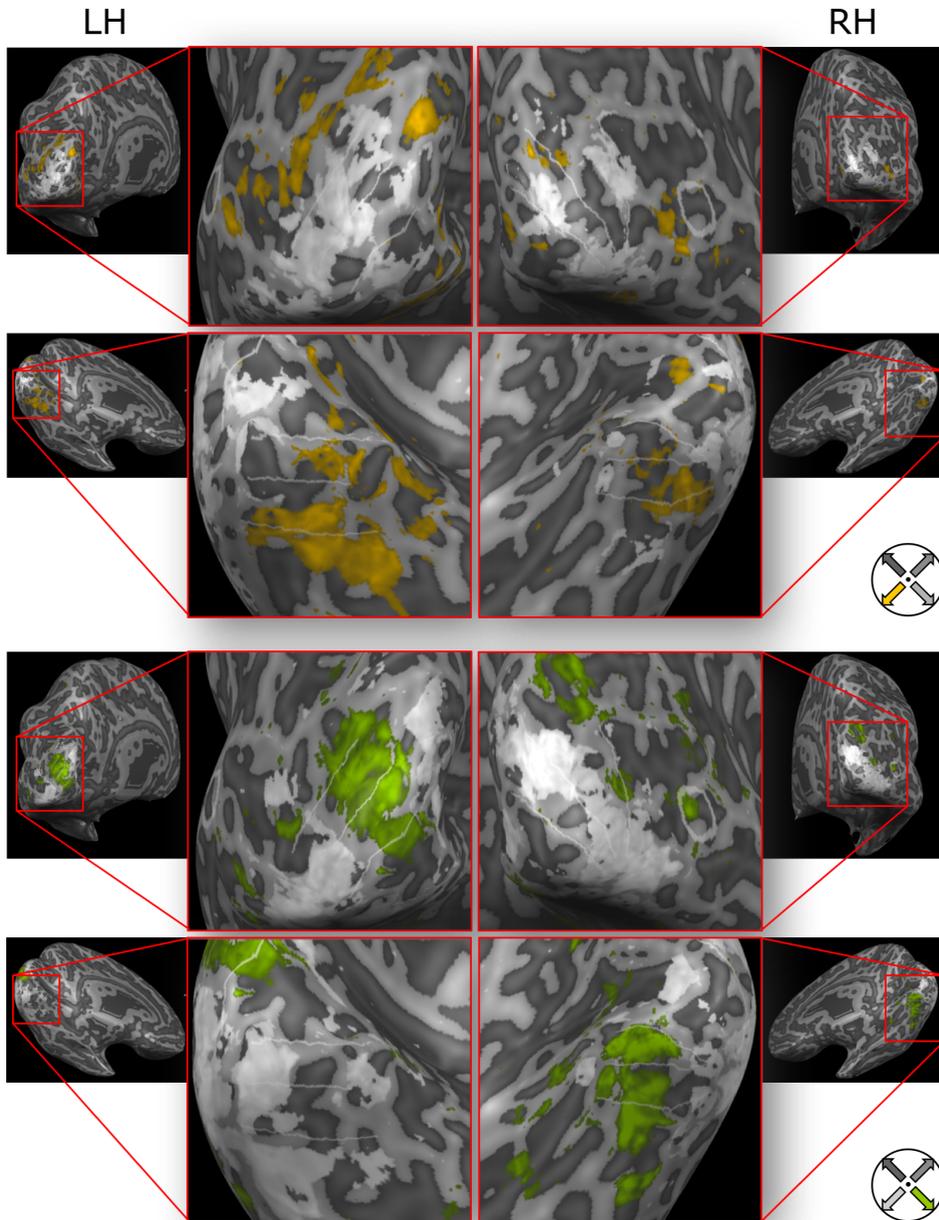


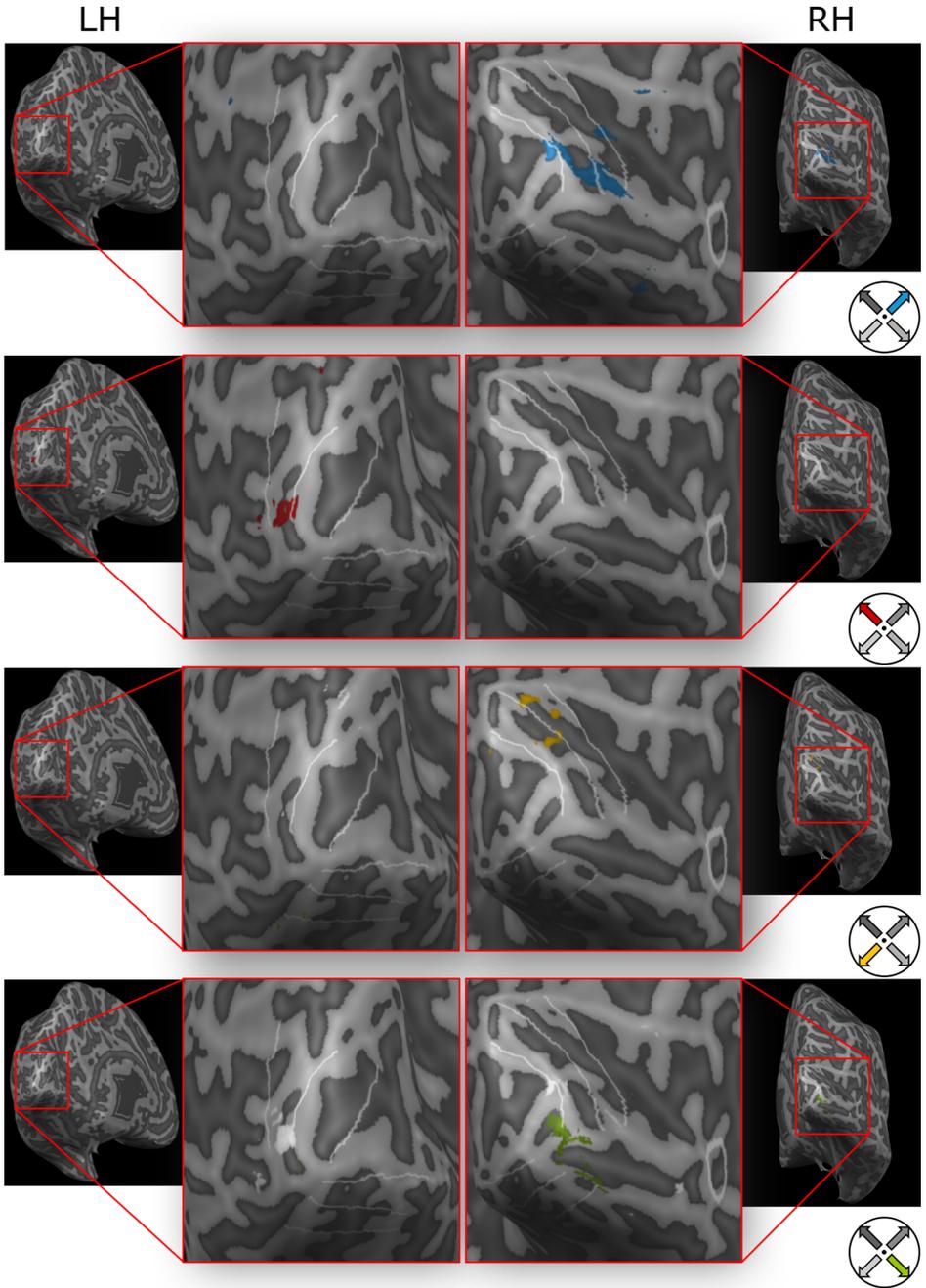
(c) Searchlight-thresholded univariate preference maps for subject 6

Figure 3.9 Searchlight-thresholded univariate preference maps for three subjects (page 92 and page 93). Preference maps for (a) subject 7, (b) subject 5, and (c) subject 6 calculated from univariate contrasts for each imagined motion direction are shown for those areas exhibiting significant effects in the searchlight map ($p < .05$ FDR corrected). Colours indicate preferred imagined motion directions and correspond to the direction wheel at the bottom right. Areas v_1 through v_4 are delineated with white lines. The MT ROI is delineated in white (white circular delineation only visible on the magnified right hemispheres of subject 7 and 5).

UNIVARIATE ANALYSIS For subject 5 and 7 single-subject GLM analyses showed significant activations for the different directions of imagined motion. Figure 3.10 shows corresponding t-maps thresholded at $q < .05$ (FDR corrected) for the same areas on the inflated cortex as in Figure 3.9a and Figure 3.9b.







(b) Univariate single-direction contrast maps for subject 5

Figure 3.10 Univariate single-direction contrast maps for two subjects (pages 94–96). T-maps (one direction versus all other directions) for (a) subject 7 (pages 94–95) and (b) subject 5 (page 96) for each imagined motion direction are shown ($q < .05$; FDR corrected). Colours correspond to the direction wheel at the bottom right. Grey/white map colours correspond to negative t-values. Areas V_1 through V_4 are delineated with white lines. The MT ROI is delineated in white (white circular delineation only visible on the magnified right hemispheres of subject 7 and 5).

EYE-TRACKING DATA No systematic eye movements associated to the different motion imagery directions were observed in subject 9 through 12 (see Figure 3.11). The test for equal median directions was not significant for any subject (S9: $P(3) = 1.159$ $p = .763$; S10: $P(3) = 2.018$ $p = .569$; S11: $P(3) = 0.471$ $p = .925$; S12: $P(3) = 0.352$ $p = .95$), that is, the saccade directions did not differ significantly between different imagined motion directions. The correlation between the direction of imagined motion and the saccade direction was also not significant for any subject (S9: $r = -.124$ $p = .37$; S10: $r = .008$ $p = .878$; S11: $r = .049$ $p = .32$; S12: $r = -.02$ $p = .666$). Furthermore, the analyses of the raw eye gaze position did neither return any significant result in the test for equal median directions (S9: $P(3) = 2.514$ $p = .473$; S10: $P(3) = 7.118$ $p = .068$; S11: $P(3) = 4.48$ $p = .214$; S12: $P(3) = 4.571$ $p = .206$) nor in the test for circular-circular correlations (S9: $r = .018$ $p = .83$; S10: $r = .047$ $p = .56$; S11: $r = .04$ $p = .626$; S12: $r = .046$ $p = .543$).

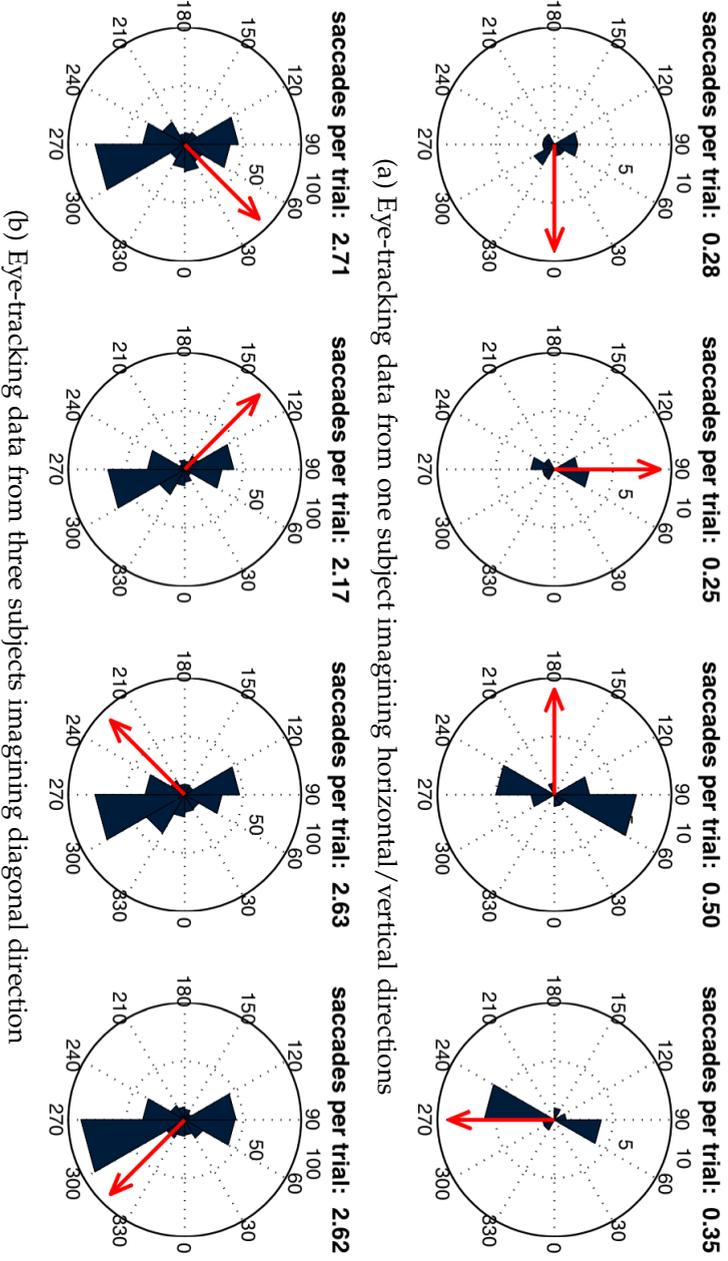


Figure 3.11 Eye-tracking data from (a) one subject imagining horizontal/vertical directions (subject 9) and (b) three subjects imagining diagonal direction (subjects 10, 11, and 12). For (a) and (b) histograms of saccade directions during trials are plotted for all four different directions of imagined motion (indicated by red arrows). The number of saccades per trial is shown above each polar plot.

3.4 DISCUSSION

Using high-resolution imaging at 7 T, we were able to successfully decode directions of imagined visual motion in fMRI data recorded from the visual system. Without any visual stimulation, subjects were able to activate their visual system so specifically that neuroimaging data recorded from areas normally activated during visual perception could predict the imagined motion direction.

The decoding accuracies ranged from non-significant results in four subjects to accuracies of up to 91.25 % (4-class classification) in a single subject – a level rarely reached even in studies decoding perceived motion.

Apart from the more global *visual motion-responsive areas* and dilated *visual motion-responsive areas* ROIs, the best classification accuracies were achieved in the V₃ and V₄ ROIs. These seem to be predominant areas to decode direction of motion during imagery. The results from the searchlight analysis complement these findings. Decoding with significant accuracies was mainly possible in the areas targeted by the ROI analyses and more pronounced in areas V₃ and V₄. More than that, the results from the searchlight analysis supported the local and potentially fine-grained origin of information in our study. Relatively local brain activation patterns – the searchlight sphere had a radius of 4.4 mm (four voxels) – predicted the imagined motion direction with substantial accuracy (up to 83.75 % in subject 7). These high decoding accuracies in single spheres might hint to a fine-grained columnar organisation of motion-selective cortical areas. Though there is an ongoing debate on whether MVPA is able to pick up patterns of brain activation at a spatial scale beyond the recorded spatial resolution (Carlson, 2014; Op de Beeck, 2010b, 2010a) and direction preferences in early visual areas seemed to have a rather coarse pattern,

High sub-categorical decoding accuracies

Localisation of decodable information

the classifiers might have picked up on a more fine-grained organisation of direction-selective columns because of a biased sampling (Haynes & Rees, 2006).

*Decoding in
hMT+*

Though the information encoded in localised MT voxels gave good classification accuracies in some subjects, optimal results were only achieved when voxels from earlier visual areas were included. This is in line with previous results revealing that the decoding of the direction of *perceived* visual motion achieves higher accuracies when analyzing data from V_1 through V_4 than from hMT+ (Kamitani & Tong, 2005). Moreover, recent work by Wang, Merriam, Freeman, and Heeger (2014) showed that the decoding of motion direction during *perception* is mainly driven by an ‘aperture-inward’ response bias producing good classification accuracies in V_1 , V_2 , and V_3 , but not hMT+. Although we recorded at a higher spatial resolution and decoded *imagined* visual motion this bias might account for the high classification accuracies in early visual areas and low classification accuracies in hMT+.

With a functional spatial resolution of 1.1 mm (iso-voxels) we are close to the resolution needed to image columnar structures. In hMT+ for instance, axis-of-motion columns are estimated to have a width of 2–2.8 mm (Zimmermann et al., 2011). A columnar organisation of hMT+, where neighboring columns prefer motion directions that are opposite (i. e., 180° difference) to form axis-of-motion columns, was proposed before (e. g., Born and Bradley, 2005). Such an organisation, however, might have actually impaired decoding in hMT+ in our study as we recorded with a spatial resolution that might just fall short of capturing single direction-of-motion columns. In order to reveal the precise organisation of direction-selective regions in early visual areas and hMT+, an even higher spatial resolution has to be reached by employing, for example, spin echo sequences that seem to be even more sensitive to local

activations than gradient echo sequences (De Martino et al., 2013; Uğurbil et al., 2003; Uludağ et al., 2009; Zimmermann et al., 2011).

When looking at the univariate statistical maps for subject 7 and 5 we observed topographic effects in mainly early and mid-level visual areas. The t-maps for each direction of imagined motion (one-versus-all) show an alignment to borders between early visual areas and to different eccentricities (foveal versus peripheral). While there is no clear organisation of this topography the significant patches overlap with the individual thresholded searchlight accuracy maps. Individual topographies might be influenced by the individual imagery style and strategy.

*Univariate
results*

We found no significant correlations between decoding accuracies and any questionnaire scores. Though the subscales of the OSIVQ showed differential correlations with decoding accuracies none of these correlations became significant. It would be important to explore such correlations in bigger sample sizes to get further insight into the possible reasons for inter-individual differences during imagery decoding. This could potentially explain the mixed results in other imagery decoding studies (Kaas, Weigelt, Roebroek, Kohler, & Muckli, 2010; Klein, Paradis, Poline, Kosslyn, & Le Bihan, 2000; Knauff, Kassubek, Mulack, & Greenlee, 2000). It has to be emphasised that the sample on which our results are based is of limited size so that we cannot rule out small or medium effects that would only be observable in bigger sample sizes.

*Inter-
individual
differences*

Due to the inherently private nature of imagery, the internal ‘stimulation’ in this experiment was hard to control, potentially inducing additional variability. However, we argue that our findings are based on visual mental imagery.

*Limitations
and outlook*

Most importantly, we decoded the direction of motion in visual areas of the cortex, thereby ruling out many confounding sources of decodable information. Furthermore, simple physiolog-

ical changes like motion artifacts or changes in breathing are too unspecific to produce the effects of the observed magnitude within the complex 4-class-classification design. Cross-modal influences on the visual system could, potentially, lead to decodable information in early visual areas. Vetter, Smith, and Muckli (2014) give an example for cross-modal influences; they showed that the content of auditory stimulation can be decoded in early visual areas. In line with this result, one could argue, that the classification results in our study are driven solely by the different auditory cues preceding the imagery trials. However, the motion imagery in our study presumably was a visuospatial imagery task with a high cognitive load (subjects reported that it would need a high cognitive engagement to maintain the continuous mental image of moving dots). In the study by Vetter et al. (2014), the addition of such a visuospatial imagery task to their original auditory stimulation eliminated most of the cross-modal influences. Therefore, we do not expect our results to be merely based on the auditory cue. Moreover, the achieved classification accuracies in our study are much higher than the classification accuracies reported in Vetter et al. (2014) and, thus, likely not only an effect of the auditory cue.

It could also be argued that the decoding of motion imagery might merely be an artifact of eye movements (see also Laeng and Teodorescu, 2002). When we analysed eye-tracking data recorded during the experiment in the fMRI scanner, there were no significant associations between the imagined motion direction and the direction of saccades. Furthermore, previous work show behavioral effects of motion imagery that cannot be explained by eye movements; Winawer, Huk, and Boroditsky (2010) had subjects imagine inward and simultaneously outward moving gratings inducing a motion aftereffect. In this experiment corresponding eye movement accounting for the observed effects, would have been

physiologically impossible. In a future study a comparison with a closed-eyes condition would be interesting as other cognitive processes might be involved when bottom-up visual input is missing entirely.

Our results could be explained by a purely attention-based modulation of visual areas. Winawer et al. (2010) argue that top-down attention modulation would only operate on feed-forward inputs to the visual cortex. However, in their study they find equal behavioral effects of mental imagery in closed- and open-eyes conditions, where in the closed-eyes condition feed-forward input is clearly missing. Even top-down modulations would still be in line with the hypothesis that imagery tasks use neural circuitry also involved in perception tasks. Our results could be caused by an internally produced stimulus that activates early visual cortex, and an attentional modulation based on the direction of imagined motion.

Finally, neural patterns underlying mental imagery have also been shown in frontal brain regions (e. g., Ganis, Thompson, and Kosslyn, 2004). Tailoring the fMRI sequence to optimally record from visual areas with a high spatial resolution, however, we were not able to record data from the frontal cortex.

Our results are of remarkable significance in single subjects and show that it is possible to reliably decode the content of complex visual imagery from neuroimaging data in single trials. Sub-categorical decoding of visual mental imagery (not ‘if’ but ‘what’) is not only relevant to advance neuroscientific knowledge, but would also enable advanced brain-computer interfaces (BCI; for reviews see Birbaumer and Cohen, 2007; Goebel, Zilverstand, and Sorger, 2010; Kübler and Neumann, 2005; Nicolas-Alonso and Gomez-Gil, 2012). In a typical BCI setup (e. g., Sorger, Reithler, Dahmen, and Goebel, 2012), the user is asked to perform a mental

*Conclusions
and
implications*

imagery tasks (e.g., imagine to move the hand versus imagine to sing a song) that corresponds to different information entities that the user wants to encode (e.g., to steer the cursor on the screen to the left versus to the right). The BCI then classifies these different categories of mental imagery tasks based on the acquired brain activation data and executes the associated intended actions (e.g., steering the cursor to the right). However, the possibility of sub-categorical decoding would allow expressing intentions in a much more natural way; the imagery of rightward visual motion (instead of imagining singing a song) could make a cursor on the screen go to the right. The recent trend towards explicit models of representation (Naselaris & Kay, 2015) would add to a closer connection between imagery content and intention. Though the use of BCIs can be trained (Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002), especially novice users rely on a close connection between imagery and translated action and would most probably benefit from such an advanced setup. In order to achieve high and stable single-trial classification accuracies that are necessary for BCIs, it will be interesting to employ spin echo sequences with an even higher spatial resolution in the future. On a sub-millimetre functional resolution columnar-level spatial separation of direction-of-motion features might become more explicitly exploitable.

REFERENCES

- Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013, August). Shared representations for working memory and mental imagery in early visual cortex. *Current biology : CB*, 23(15), 1427–31. doi:[10.1016/j.cub.2013.05.065](https://doi.org/10.1016/j.cub.2013.05.065)
- Albright, T. D. (1984, December). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of neurophysiology*, 52(6), 1106–30. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6520628>
- Albright, T. D., Desimone, R., & Gross, C. G. (1984, January). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of neurophysiology*, 51(1), 16–31. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6693933>
- Andersson, J. L. R., Skare, S., & Ashburner, J. (2003, October). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, 20(2), 870–88. doi:[10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7)
- Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *J Stat Softw*, 31(10).
- Birbaumer, N. & Cohen, L. G. (2007, March). Brain-computer interfaces: communication and restoration of movement in paralysis. *The Journal of physiology*, 579(Pt 3), 621–36. doi:[10.1113/jphysiol.2006.125633](https://doi.org/10.1113/jphysiol.2006.125633)
- Blazhenkova, O. & Kozhevnikov, M. (2009, July). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, 23(5), 638–663. doi:[10.1002/acp.1473](https://doi.org/10.1002/acp.1473)
- Bonhoeffer, T. & Grinvald, A. (1991, October). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature*, 353(6343), 429–31. doi:[10.1038/353429a0](https://doi.org/10.1038/353429a0)
- Born, R. T. & Bradley, D. C. (2005, January). Structure and function of visual area MT. *Annual review of neuroscience*, 28, 157–89. doi:[10.1146/annurev.neuro.26.041002.131052](https://doi.org/10.1146/annurev.neuro.26.041002.131052)
- Carlson, T. A. (2014, June). Orientation decoding in human visual cortex: new insights from an unbiased perspective. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 34(24), 8373–83. doi:[10.1523/JNEUROSCI.0548-14.2014](https://doi.org/10.1523/JNEUROSCI.0548-14.2014)

- Chang, C.-C. & Lin, C.-J. (2011, April). Libsvm. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–27. doi:[10.1145/1961189.1961199](https://doi.org/10.1145/1961189.1961199)
- Cheng, K., Fujita, H., Kanno, I., Miura, S., & Tanaka, K. (1995, July). Human cortical regions activated by wide-field visual motion: an H₂(¹⁵O) PET study. *Journal of neurophysiology*, 74(1), 413–27. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7472342>
- Cheng, K., Waggoner, R. A., & Tanaka, K. (2001, October). Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. *Neuron*, 32(2), 359–74. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11684004>
- Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012, February). Imagery and perception share cortical representations of content and location. *Cerebral cortex (New York, N.Y. : 1991)*, 22(2), 372–80. doi:[10.1093/cercor/bhr106](https://doi.org/10.1093/cercor/bhr106)
- De Martino, F., Zimmermann, J., Muckli, L., Uğurbil, K., Yacoub, E., & Goebel, R. (2013, January). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. *PLoS one*, 8(3), e60514. doi:[10.1371/journal.pone.0060514](https://doi.org/10.1371/journal.pone.0060514)
- Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: Analyses and review. *Journal of Experimental Psychology: General*, 140(4), 552–572. doi:[10.1037/a0023885](https://doi.org/10.1037/a0023885)
- Diogo, A. C. M., Soares, J. G. M., Koulakov, A., Albright, T. D., & Gattass, R. (2003, May). Electrophysiological imaging of functional architecture in the cortical middle temporal visual area of Cebus apella monkey. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 23(9), 3881–98. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12736358>
- Dubner, R. & Zeki, S. M. (1971, December). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain research*, 35(2), 528–32. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/5002708>
- Dumoulin, S. O. & Wandell, B. A. (2008, January). Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2), 647–60. doi:[10.1016/j.neuroimage.2007.09.034](https://doi.org/10.1016/j.neuroimage.2007.09.034)
- Dupont, P., Orban, G. A., De Bruyn, B., Verbruggen, A., & Mortelmans, L. (1994, September). Many areas in the human brain respond to visual motion.

- Journal of neurophysiology*, 72(3), 1420–4. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7807222>
- Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *The Journal of Neuroscience*, 33(15), 6516–6523. doi:10.1523/JNEUROSCI.5732-12.2013
- Engbert, R. & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences of the United States of America*, 103(18), 7192–7. doi:10.1073/pnas.0509557103
- Ferguson, E. S. (1977, August). The Mind's Eye: Nonverbal Thought in Technology. *Science (New York, N.Y.)* 197(4306), 827–36. doi:10.1126/science.197.4306.827
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004, July). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Brain research. Cognitive brain research*, 20(2), 226–41. doi:10.1016/j.cogbrainres.2004.02.012
- Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of Functional Image Analysis Contest (FIAC) data with BrainVoyager QX: From single-subject to cortically aligned group General Linear Model analysis and self-organizing group Independent Component Analysis. *Human Brain Mapping*, 27(5), 392–401. doi:10.1002/hbm.20249
- Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998, May). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *The European journal of neuroscience*, 10(5), 1563–73. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9751129>
- Goebel, R., Zilverstand, A., & Sorger, B. (2010, August). Real-time fMRI-based brain-computer interfacing for neurofeedback therapy and compensation of lost motor functions. *Imaging in Medicine*, 2(4), 407–415. doi:10.2217/iim.10.35
- Goodyear, B. G. & Menon, R. S. (2001, December). Brief visual stimulation allows mapping of ocular dominance in visual cortex using fMRI. *Human brain mapping*, 14(4), 210–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11668652>
- Harrison, S. A. & Tong, F. (2009, April). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–5. doi:10.1038/nature07832

- Haynes, J.-D. & Rees, G. (2006, July). Decoding mental states from brain activity in humans. *Nature reviews. Neuroscience*, 7(7), 523–34. doi:[10.1038/nrn1931](https://doi.org/10.1038/nrn1931)
- Hubel, D. H. & Wiesel, T. N. (1968, January). Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 67, 40–113. doi:[10.1113/jphysiol.1968.sp008455](https://doi.org/10.1113/jphysiol.1968.sp008455)
- Johnson, M. R. & Johnson, M. K. (2014, January). Decoding individual natural scene representations during perception and imagery. *Frontiers in human neuroscience*, 8(February), 59. doi:[10.3389/fnhum.2014.00059](https://doi.org/10.3389/fnhum.2014.00059)
- Kaas, A., Weigelt, S., Roebroek, A., Kohler, A., & Muckli, L. (2010, January). Imagery of a moving object: the role of occipital cortex and human MT/V5+. *NeuroImage*, 49(1), 794–804. doi:[10.1016/j.neuroimage.2009.07.055](https://doi.org/10.1016/j.neuroimage.2009.07.055)
- Kamitani, Y. & Tong, F. (2005, May). Decoding the visual and subjective contents of the human brain. *Nature neuroscience*, 8(5), 679–85. doi:[10.1038/nn1444](https://doi.org/10.1038/nn1444)
- Kamitani, Y. & Tong, F. (2006, June). Decoding seen and attended motion directions from activity in the human visual cortex. *Current biology : CB*, 16(11), 1096–102. doi:[10.1016/j.cub.2006.04.003](https://doi.org/10.1016/j.cub.2006.04.003)
- Klein, I., Paradis, A. L., Poline, J.-B., Kosslyn, S. M., & Le Bihan, D. (2000, January). Transient activity in the human calcarine cortex during visual-mental imagery: an event-related fMRI study. *Journal of cognitive neuroscience*, 12 Suppl 2, 15–23. doi:[10.1162/089892900564037](https://doi.org/10.1162/089892900564037)
- Knauff, M., Kassubek, J., Mulack, T., & Greenlee, M. W. (2000, December). Cortical activation evoked by visual mental imagery as measured by fMRI. *NeuroReport*, 11(18), 3957–3962. doi:[10.1097/00001756-200012180-00011](https://doi.org/10.1097/00001756-200012180-00011)
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001, September). Neural foundations of imagery. *Nature reviews. Neuroscience*, 2(9), 635–42. doi:[10.1038/35090055](https://doi.org/10.1038/35090055)
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006, March). *The Case for Mental Imagery*. Oxford University Press. doi:[10.1093/acprof:oso/9780195179088.001.0001](https://doi.org/10.1093/acprof:oso/9780195179088.001.0001)
- Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002, March). Revising the Visualizer-Verbalizer Dimension: Evidence for Two Types of Visualizers. *Cognition and Instruction*, 20(1), 47–77. doi:[10.1207/S1532690XCI2001_3](https://doi.org/10.1207/S1532690XCI2001_3)
- Kozhevnikov, M., Kosslyn, S. M., & Shephard, J. (2005, June). Spatial versus object visualizers: a new characterization of visual cognitive style. *Memory & cognition*, 33(4), 710–26. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16248335>

- Kozhevnikov, M., Kozhevnikov, M., Yu, C. J., & Blazhenkova, O. (2013, June). Creativity, visualization abilities, and visual cognitive style. *The British journal of educational psychology*, 83(Pt 2), 196–209. doi:[10.1111/bjep.12013](https://doi.org/10.1111/bjep.12013)
- Kriegeskorte, N. & Goebel, R. (2001). An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical mr volumes. *NeuroImage*, 14(2), 329–346. doi:[10.1006/nimg.2001.0831](https://doi.org/10.1006/nimg.2001.0831)
- Kriegeskorte, N., Goebel, R., & Bandettini, P. A. (2006, March). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, 103(10), 3863–8. doi:[10.1073/pnas.0600244103](https://doi.org/10.1073/pnas.0600244103)
- Kübler, A. & Neumann, N. (2005, January). Brain-computer interfaces—the key for the conscious brain locked into a paralyzed body. *Progress in brain research*, 150, 513–25. doi:[10.1016/S0079-6123\(05\)50035-9](https://doi.org/10.1016/S0079-6123(05)50035-9)
- Laeng, B. & Teodorescu, D.-S. (2002, March). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, 26(2), 207–231. doi:[10.1207/s15516709cog2602_3](https://doi.org/10.1207/s15516709cog2602_3)
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2012, February). Disentangling visual imagery and perception of real-world objects. *NeuroImage*, 59(4), 4064–73. doi:[10.1016/j.neuroimage.2011.10.055](https://doi.org/10.1016/j.neuroimage.2011.10.055)
- Marks, D. F. (1973, February). Visual imagery differences in the recall of pictures. *British journal of psychology (London, England : 1953)*, 64(1), 17–24. doi:[10.1111/j.2044-8295.1973.tb01322.x](https://doi.org/10.1111/j.2044-8295.1973.tb01322.x)
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., . . . Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 356(1412), 1293–1322. doi:[10.1098/rstb.2001.0915](https://doi.org/10.1098/rstb.2001.0915)
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., . . . Mazoyer, B. (2001). A Four-Dimensional Probabilistic Atlas of the Human Brain. *Journal of the American Medical Informatics Association*, 8(5), 401–430. doi:[10.1136/jamia.2001.0080401](https://doi.org/10.1136/jamia.2001.0080401)
- McLean, J. & Palmer, L. A. (1989, January). Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Research*, 29(6), 675–679. doi:[10.1016/0042-6989\(89\)90029-1](https://doi.org/10.1016/0042-6989(89)90029-1)
- Menon, R. S., Ogawa, S., Strupp, J. P., & Uğurbil, K. (1997, May). Ocular dominance in human V1 demonstrated by functional magnetic resonance

- imaging. *Journal of neurophysiology*, 77(5), 2780–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9163392>
- Miller, A. I. (1996). *Insights of Genius*. New York, NY: Springer New York. doi:10.1007/978-1-4612-2388-7
- Moshtagh, N. (2005). Minimum volume enclosing ellipsoids. *Convex Optimization*, 1–9. doi:10.1.1.116.7691
- Motes, M. A., Malach, R., & Kozhevnikov, M. (2008, November). Object-processing neural efficiency differentiates object from spatial visualizers. *Neuroreport*, 19(17), 1727–31. doi:10.1097/WNR.0bo13e328317f3e2
- Mountcastle, V. B. (1957, July). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of neurophysiology*, 20(4), 408–34. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/13439410>
- Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978, October). Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *The Journal of physiology*, 283, 101–20.
- Naselaris, T. & Kay, K. N. (2015). Resolving Ambiguities of MVPA Using Explicit Models of Representation. *Trends in cognitive sciences*, 19(10), 551–4. doi:10.1016/j.tics.2015.07.005
- Nicolas-Alonso, L. F. & Gomez-Gil, J. (2012, January). Brain computer interfaces, a review. *Sensors (Basel, Switzerland)*, 12(2), 1211–79. doi:10.3390/s120201211
- Op de Beeck, H. P. (2010a, February). Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage*, 49(3), 1943–8. doi:10.1016/j.neuroimage.2009.02.047
- Op de Beeck, H. P. (2010b, April). Probing the mysterious underpinnings of multi-voxel fMRI analyses. *NeuroImage*, 50(2), 567–71. doi:10.1016/j.neuroimage.2009.12.072
- Peirce, J. W. (2007, May). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2), 8–13. doi:10.1016/j.jneumeth.2006.11.017
- Pereira, F., Mitchell, T., & Botvinick, M. (2009, March). Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage*, 45(1 Suppl), S199–209. doi:10.1016/j.neuroimage.2008.11.007
- Poltrock, S. & Brown, P. (1984). Individual differences in visual imagery and spatial ability. *Intelligence*, 138, 93–138. doi:10.1016/0160-2896(84)90019-9

- Reddy, L., Tsuchiya, N., & Serre, T. (2010, April). Reading the mind's eye: decoding category information during mental imagery. *NeuroImage*, 50(2), 818–25. doi:[10.1016/j.neuroimage.2009.11.084](https://doi.org/10.1016/j.neuroimage.2009.11.084)
- Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N. K., & Uğurbil, K. (2007, April). Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. *NeuroImage*, 35(2), 539–52. doi:[10.1016/j.neuroimage.2006.12.030](https://doi.org/10.1016/j.neuroimage.2006.12.030)
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., . . . Matthews, P. M. (2004, January). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1, S208–19. doi:[10.1016/j.neuroimage.2004.07.051](https://doi.org/10.1016/j.neuroimage.2004.07.051)
- Sorger, B., Reithler, J., Dahmen, B., & Goebel, R. (2012, July). A real-time fMRI-based spelling device immediately enabling robust motor-independent communication. *Current biology : CB*, 22(14), 1333–8. doi:[10.1016/j.cub.2012.05.022](https://doi.org/10.1016/j.cub.2012.05.022)
- Stokes, M., Saraiva, A., Rohenkohl, G., & Nobre, A. C. (2011, June). Imagery for shapes activates position-invariant representations in human visual cortex. *NeuroImage*, 56(3), 1540–5. doi:[10.1016/j.neuroimage.2011.02.071](https://doi.org/10.1016/j.neuroimage.2011.02.071)
- Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009, February). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(5), 1565–72. doi:[10.1523/JNEUROSCI.4657-08.2009](https://doi.org/10.1523/JNEUROSCI.4657-08.2009)
- Swisher, J. D., Gatenby, J. C., Gore, J. C., Wolfe, B. a., Moon, C.-H., Kim, S.-G., & Tong, F. (2010, January). Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(1), 325–30. doi:[10.1523/JNEUROSCI.4811-09.2010](https://doi.org/10.1523/JNEUROSCI.4811-09.2010)
- Tanaka, K. (1996, January). Inferotemporal cortex and object vision. *Annual review of neuroscience*, 19, 109–39. doi:[10.1146/annurev.ne.19.030196.000545](https://doi.org/10.1146/annurev.ne.19.030196.000545)
- Tootell, R. B. H., Mendola, J. D., Hadjikhani, N. K., Ledden, P. J., Liu, a. K., Reppas, J. B., . . . Dale, A. M. (1997, September). Functional analysis of V3A and related areas in human visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 17(18), 7060–78. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9278542>
- Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., . . . Belliveau, J. W. (1995, April). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal*

- of neuroscience : the official journal of the Society for Neuroscience*, 15(4), 3215–30. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7722658>
- Uğurbil, K., Adriany, G., Andersen, P., Chen, W., Garwood, M., Gruetter, R., ... Zhu, X.-H. (2003, December). Ultrahigh field magnetic resonance imaging and spectroscopy. *Magnetic Resonance Imaging*, 21(10), 1263–1281. doi:10.1016/j.mri.2003.08.027
- Uludağ, K., Müller-Bierl, B., & Uğurbil, K. (2009, October). An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *NeuroImage*, 48(1), 150–65. doi:10.1016/j.neuroimage.2009.05.051
- Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1997). The kinetic occipital (KO) region in man: an fMRI study. *Cerebral cortex (New York, N.Y. : 1991)*, 7(7), 690–701. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9373023>
- Vaughan, J. T., Garwood, M., Collins, C. M., Liu, W., DelaBarre, L., Adriany, G., ... Uğurbil, K. (2001, July). 7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, 46(1), 24–30. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11443707>
- Vetter, P., Smith, F. W., & Muckli, L. (2014, June). Decoding sound and imagery content in early visual cortex. *Current biology : CB*, 24(11), 1256–62. doi:10.1016/j.cub.2014.04.020
- Wang, H. X., Merriam, E. P., Freeman, J., & Heeger, D. J. (2014, September). Motion direction biases and decoding in human visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 34(37), 12601–15. doi:10.1523/JNEUROSCI.1034-14.2014
- Winawer, J., Huk, A. C., & Boroditsky, L. (2010, February). A motion aftereffect from visual imagery of motion. *Cognition*, 114(2), 276–84. doi:10.1016/j.cognition.2009.09.010
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002, June). Brain-computer interfaces for communication and control. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 113(6), 767–91. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12048038>

- Wraga, M., Thompson, W. L., Alpert, N. M., & Kosslyn, S. M. (2003, July). Implicit transfer of motor strategies in mental rotation. *Brain and Cognition*, 52(2), 135–143. doi:[10.1016/S0278-2626\(03\)00033-2](https://doi.org/10.1016/S0278-2626(03)00033-2)
- Yacoub, E., Harel, N., & Uğurbil, K. (2008, July). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(30), 10607–12. doi:[10.1073/pnas.0804110105](https://doi.org/10.1073/pnas.0804110105)
- Yacoub, E., Shmuel, A., Logothetis, N. K., & Uğurbil, K. (2007, October). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *NeuroImage*, 37(4), 1161–77. doi:[10.1016/j.neuroimage.2007.05.020](https://doi.org/10.1016/j.neuroimage.2007.05.020)
- Zeman, A., Dewar, M., & Della Sala, S. (2015, December). Lives without imagery – Congenital aphantasia. *Cortex*, 73, 378–380. doi:[10.1016/j.cortex.2015.05.019](https://doi.org/10.1016/j.cortex.2015.05.019)
- Zimmermann, J., Goebel, R., De Martino, F., Van de Moortele, P.-F., Feinberg, D. A., Adriany, G., . . . Yacoub, E. (2011, December). Mapping the Organization of Axis of Motion Selective Features in Human Area MT Using High-Field fMRI. *PLoS ONE*, 6(12), e28716. doi:[10.1371/journal.pone.0028716](https://doi.org/10.1371/journal.pone.0028716)

Chapter 4

Reconstructing and decoding imagined letters from the visual system using ultra-high field fMRI

based on: Emmerling, T. C., Senden, M., Frost, M. A., & Goebel, R. (*in preparation*). Reconstructing and decoding imagined letters from the visual system using ultra-high field fMRI.

We would like to thank Franziska Emmerling and Carmine Gnolo for their help.

ABSTRACT

Research employing high-resolution functional imaging data has succeeded in reconstructing visual field images of perceived content from evoked activity in human visual cortex. A more challenging goal of cognitive neuroscience is to unravel not only the neural underpinnings of visual perception, but those of visual mental imagery, a process which is considered one of the main sources of human cognitive complexity. While recent work has demonstrated functional retinotopic activation of early visual areas during visual mental imagery, it is not yet specified exactly how a mental image of a stimulus is composed and how it compares to its perceived counterpart or an actually presented stimulus. In the current study, we demonstrate for the first time that it is possible to reconstruct visual field images which carry recognisable content of imagined letter shapes, even on a single trial level. Using submillimetre resolution functional magnetic resonance imaging (fMRI) data of two subjects and individually computed population receptive field mappings, we reconstructed the visual field during perception and imagery of four different letter shapes and found a remarkable similarity to the presented letter shapes, especially in the earliest visual areas. We demonstrate significant overlap between reconstructed visual field images of perception data, imagery data, and the actually presented stimuli. Furthermore, we successfully classified between different letter features (edgy versus curvy) using multi-voxel pattern data from visual word form area (VWFA). This proof of concept is the first step towards more elaborate studies on visual field image reconstruction of imagined everyday stimuli and towards the development of more natural and directly content-based brain-computer-interfaces.

4.1 INTRODUCTION

Visual perception and visual mental imagery share neural circuitry in the human brain (Kosslyn & Thompson, 2003; Kosslyn, Thompson, & Ganis, 2006; Pearson, Naselaris, Holmes, & Kosslyn, 2015). Over the last decades, several studies have shown that visual mental imagery activates cortical networks that are also activated during corresponding perceptual tasks (Ganis, Thompson, & Kosslyn, 2004; Goebel, Khorrám-Sefat, Muckli, Hacker, & Singer, 1998; Ishai, Ungerleider, & Haxby, 2000; Kosslyn, Thompson, & Alpert, 1997; Mechelli, Price, Friston, & Ishai, 2004; O'Craven & Kanwisher, 2000). Applying different forms of machine learning approaches to functional magnetic resonance imaging (fMRI) data enabled the decoding of imagery *content* regarding visual mental imagery of orientations (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Harrison & Tong, 2009), motion (Emmerling, Zimmermann, Sorger, Frost, & Goebel, 2016), objects (Cichy, Heinzle, & Haynes, 2012; Lee, Kravitz, & Baker, 2012; Reddy, Tsuchiya, & Serre, 2010), shapes (Stokes, Saraiva, Rohenkohl, & Nobre, 2011; Stokes, Thompson, Cusack, & Duncan, 2009), and scenes (M. R. Johnson & Johnson, 2014; Naselaris, Olman, Stansbury, Uğurbil, & Gallant, 2014).

During many visual mental imagery tasks, a mental image similar to its perceptual counterpart is formed (Kosslyn, Thompson, Wraga, & Alpert, 2001). This mental image is of an inherently private nature. Recent advancements in the reconstruction of *perceived* visual stimuli from fMRI data (Miyawaki et al., 2008; Schoenmakers, Barth, Heskes, & van Gerven, 2013; Thirion et al., 2006) – i. e. a visualisation of what a subject saw based on their brain activations – pose the question of whether reconstruction rather than mere decoding is also possible for a mental image. Studies that

reconstruct visual perception based on neuroimaging data leveraged the retinotopic organisation of early visual areas (Sereno et al., 1995) and fit sophisticated encoding models to the visual tuning parameters of these areas. The method of population receptive field (pRF) mapping Dumoulin and Wandell, 2008 enabled a straightforward and fast estimation of such parameters for each fMRI voxel recorded from the visual system. These methodological advances, which enabled the reconstruction of perceived images from fMRI data, have the potential to be taken one step further: They can be applied to brain activation elicited by merely imagined visual images. Previous studies (Klein et al., 2004; Slotnick, Thompson, & Kosslyn, 2005) already demonstrated functionally specific retinotopic activations during imagery. Beyond these results, a visual reconstruction of a complex mental image from neuroimaging data would be conclusive evidence that mental imagery can activate visual cortex in an extensive and precise manner. A mental image reconstruction could also demonstrate possible distortions in such an image that have been proposed before (Naselaris & Lynam, 2014). Finally, the comparison between cortical representations of perception and imagery (Bartolomeo, 2002) can be investigated conservatively on the level of reconstructed visual fields from the two activated patterns.

Letter imagery lends itself well to a mental image reconstruction experiment, as proficient readers are very familiar with letter shapes. It is easy for subjects to imagine the rather complex mental visual image of a letter while this mental image is not yet too fine-grained to be reconstructed given current neuroimaging and retinotopic resolutions (Schoenmakers et al., 2013). In contrast to an imagined artwork (Naselaris et al., 2014) or natural scenes (M. R. Johnson & Johnson, 2014), letters are highly overlearned and abstract stimuli with simple geometric properties enabling compa-

rably well-controlled imagery conditions. Furthermore, brain activity during letter imagery might also involve specialised areas of the visual system that are activated during perception of letter shapes including mid-level visual word form area (VWFA; McCandliss, Cohen, and Dehaene, 2003). Brain activity patterns in these visual areas might reflect subtle feature-level differences between different imagined letter shapes and, thereby, improve decoding results.

In this study we set out to reconstruct and decode imagined letters from submillimetre resolution fMRI data recorded in human visual cortex using individual pRF mappings. We investigated whether visual field image (VFI) reconstructions of four letter shapes (perceived and imagined) were possible and in how far VFIs, elicited by perception versus imagery, overlapped with each other and the actual presented stimulus. Furthermore, we tested the success of VFI reconstructions in single imagery trials and classified these reconstructions using perception generated VFIs as well as the actual presented stimuli as reference. Finally, we complete our proof of concept by examining the involvement of higher visual areas and classifying between letters with a predominantly 'edgy' or 'curvy' morphology using brain activity patterns in VWFA.

4.2 METHODS

SUBJECTS Two subjects (one female) with normal vision took part in this study. All subjects were experienced in participating in high field fMRI experiments, gave written informed consent and were paid for participation. All procedures were conducted with

approval from the local Ethical Committee of the Faculty of Psychology and Neuroscience at Maastricht University.

TRAINING SESSION The subject completed three training sessions prior to the actual experiments to practice the controlled imagery of visual letters. Each training session lasted ca. 45 minutes and was scheduled in the week before the scanning session. Before the first training session, the subject filled in the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973) and the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ; Blazhenkova and Kozhevnikov, 2009). These questionnaires measure the subjective clearness and vividness of imagined objects and cognitive styles during mental imagery, respectively.

TRAINING TASK In each training trial, the subject saw one of four white letters ('H', 'T', 'S', or 'C') in a square white guide box ($8^\circ \times 8^\circ$ visual angle) on grey background and a red fixation dot in the center of the screen (see [Figure 4.1](#)). With the onset of the visual stimulation the subject heard a pattern of three low tones (note C5) and one high tone (note G5) that lasted 1000 ms. This tone pattern was associated with the visually presented letter. After 3000 ms the letter started to fade out and was completely gone 5000 ms after trial onset. The fixation dot then turned orange and the subject was instructed to maintain a vivid image of the presented letter. After an 18 s imagery period, the fixation dot turned white and probing started. With an inter-probe-interval of 1500 ms (jittered by ± 200 ms) three white probe dots appeared within the guide box. These dots were located within the letter shape or outside of the letter shape (however, always within the guide box). The subject was instructed to indicate by button press whether a probe was located inside or outside the imagined letter shape. Depending on

the response, the fixation dot turned red (incorrect) or green (correct) before turning white again as soon as the next probe was shown. The positions of the probe dots were randomly chosen so that they had a minimum distance of 0.16° and a maximum distance of 0.32° of visual angle from the edges of the letter (and the guide box), both for inside and outside probes. This ensured similar task difficulty across trials. A resting phase of 3000 ms or 6000 ms followed the three probes. At the beginning of a training run all four letters were presented for 3000 ms each, alongside the associated tone pattern (reference phase). During one training run, the subject completed 16 pseudo-randomly presented trials. In each training session the subject completed two training runs during which reference letters were presented in each trial (described above) and two training runs without visual presentation (so that the tone pattern was the only cue for a letter). The subject was instructed to fixate the fixation dot during the whole run. After the training session the subject verbally reported the imagery strategies that they had used to the experimenter. The letter-tone associations were randomly assigned for each subject.

EXPERIMENTAL TASK The experimental task was similar to the training task with changes to the probing phase and the timing of the trial phase. After the reference phase in the beginning of each run, there was no visual stimulation other than the fixation dot and the guide box. Imagery phases started when the subject heard the tone pattern and the fixation dot turned orange. Imagery phases lasted 6 s. The subject was instructed to imagine the letter associated to the tone pattern as vividly and accurately as possible. The guide box aided the participant by acting as a reference for the physical dimensions of the letter. The resting phases that followed each imagery phase lasted 9 s or 12 s. There was no prob-

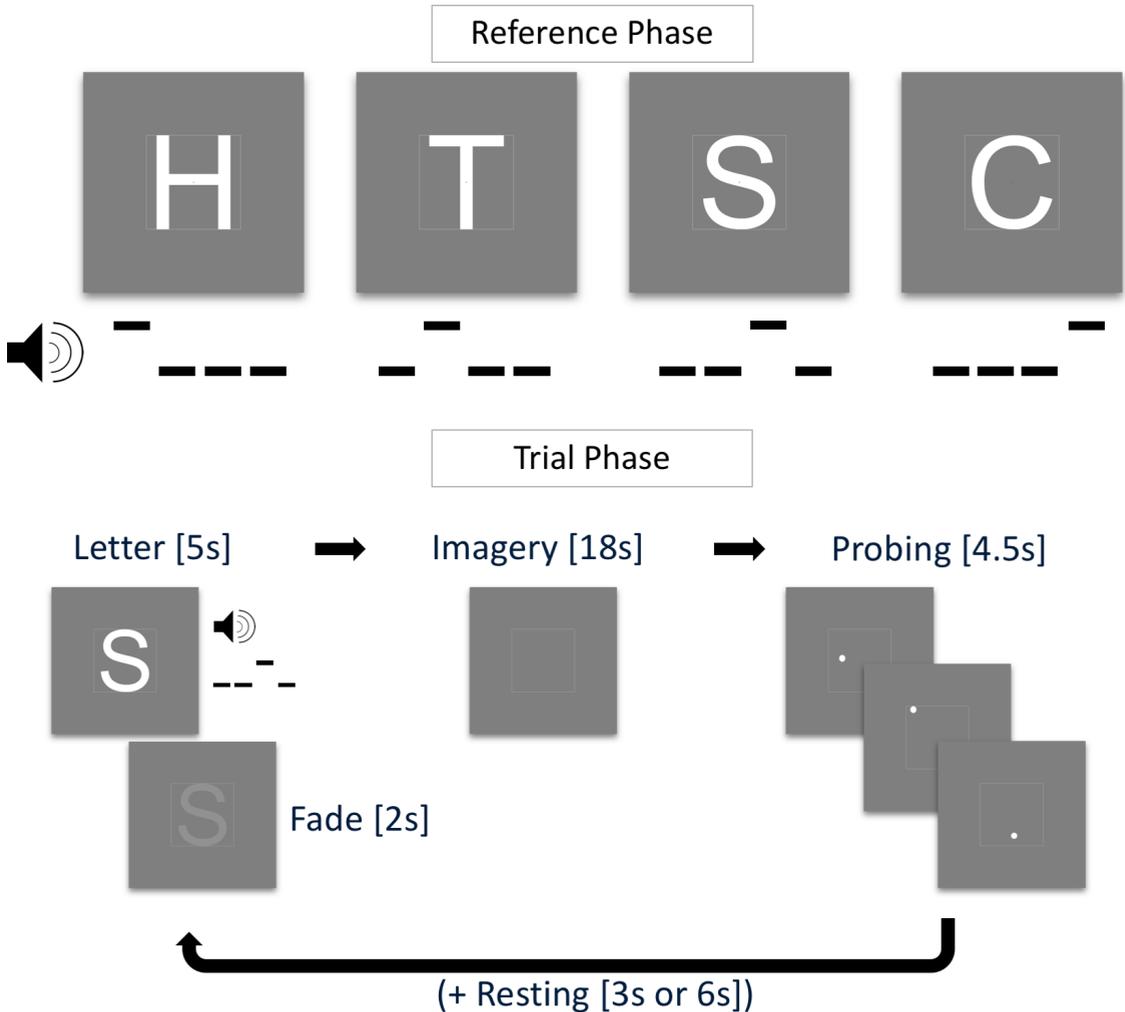


Figure 4.1 Training task. Visual and auditory stimulation are visualised schematically for reference phase (top box) and trial phase (bottom box). In the reference phase, the four letters 'H', 'T', 'S', and 'C' are presented together with their respectively associated tone pattern of high and low tones. Here, this association is shown for subject S₁. In the trial phase, the tone pattern is played and the letter is shown for 5 s (fading out after 3 s) followed by an imagery period of 18 s, a probing period of three times 1.5 s, and finally a resting period of 3 s or 6 s. The fixation dot and the white guide box stay on the screen for the whole run. Note: Probe dots are magnified for better visualisation.

ing phase in normal trials. In each experimental run there were 32 normal trials and two additional catch trials in which there was a probing phase that had four probes each. There was no visual feedback for the responses in the probing phase (the fixation dot remained white). Data from the catch trials were not included into the analysis.

PERCEPTION RUNS To measure brain activation patterns in visual areas during the perception of the letters used in the imagery runs we recorded one perception run during the scanning session. The four letters were visually presented using the same trial timing parameters as in the experimental runs. There was no reference or probing phases. Letters were presented for the duration of the imagery phase (6 s) and their shape was filled with a flickering checkerboard pattern (10 Hz). No tone patterns were played during the perception run.

STIMULATION EQUIPMENT The visual and auditory stimulation was controlled with PsychoPy (version 1.83.03; Peirce, 2007). Visual stimuli were projected on a frosted screen at the top end of the scanner table by means of an LCD projector (Panasonic, No PT-EZ57OEL; Newark, NJ, USA). Auditory stimulation was presented using MR-compatible insert earphones (Sensimetrics, Model S14; Malden, MA, USA). Responses to the probes were recorded with MR-compatible button boxes (Current Designs, 8-button response device, HHSC-2x4-C; Philadelphia, USA).

SCANNING SESSION At the beginning of the scanning session we recorded a pRF retinotopy mapping (Dumoulin & Wandell, 2008) with flickering checkerboard bars in twelve different positions and eight different angles lasting in total about

16 minutes. Subsequently, we recorded two experimental imagery runs (each about 11 minutes) followed by one perception run (about 9 minutes) and two additional experimental imagery runs. In the end of the scanning session we recorded one run (subject S2) or two runs (subject S1) of a VWFA localizer (Stigliani, Weiner, & Grill-Spector, 2015).

MRI ACQUISITION We recorded anatomical and functional images with a Siemens Magnetom 7T scanner (Siemens; Erlangen, Germany) and a 32-channel head-coil (Nova Medical Inc.; Wilmington, MA, USA). Before the functional scans we used a T1-weighted magnetisation prepared rapid acquisition gradient echo (3D-MP2RAGE; Marques et al., 2010) sequence (240 sagittal slices, matrix = 320×320 , voxel size = $0.7 \times 0.7 \times 0.7 \text{ mm}^3$, first inversion time TI1 = 900 ms, second inversion time TI2 = 2750 ms, echo time (TE) = 2.46 ms, repetition time (TR) = 5000 ms, first nominal flip angle = 5° , second nominal flip angle = 3°) to acquire an anatomical dataset. For all functional runs we acquired high-resolution gradient echo (T2* weighted) echo-planar imaging (EPI; Moeller et al., 2010) data (TE = 26 ms, TR = 3000 ms, generalised auto-calibrating partially parallel acquisitions (GRAPPA) factor = 3, multi-band factor = 2, nominal flip angle = 55° , number of slices = 82, matrix = 186×186 , voxel size = $0.8 \times 0.8 \times 0.8 \text{ mm}^3$). The field-of-view covered occipital, parietal, and temporal areas, including inferior posterior regions like VWFA. Additionally, before the first functional scan we recorded five functional volumes with opposed phase encoding directions to correct for EPI distortions that occur at higher field strengths.

IMAGE DATA PREPROCESSING We analysed anatomical and functional images using BrainVoyager 20 (version 20.0; Brain In-

novation; Maastricht, The Netherlands), PyMVPA (version 2.3; Hanke et al., 2009), and custom code in MATLAB (version 2015a; The Mathworks Inc.; Natick, MA, USA). We interpolated anatomical images to a nominal resolution of 0.8 mm isovoxel size matching the functional images resolution. In the anatomical images, the grey/white matter boundary was detected and segmented using the advanced automatic segmentation tools of BrainVoyager 20 that are optimised for high-field MRI data. A region-growing approach analysed local intensity histograms, corrected topological errors of the segmented grey/white matter border and finally reconstructed meshes of the cortical surfaces (Goebel, Esposito, & Formisano, 2006; Kriegeskorte & Goebel, 2001). The functional images were corrected for motion artefacts using the 3D rigid body motion correction algorithm implemented in BrainVoyager 20 and all functional runs were aligned to the first volume of the first functional run. We corrected the EPI distortions using a method similar to that described in Andersson, Skare, and Ashburner (2003) and the ‘topup’ tool implemented in FSL (Smith et al., 2004). The pairs of reversed phase encoding images recorded in the beginning of the scanning session were used to estimate the susceptibility-induced off-resonance field and, then, correct the distortions in the remaining functional runs. After this correction, functional data were high-pass filtered using a general linear model (GLM) Fourier basis set of three (all functional runs apart the VWFA localizer) or four (VWFA localizer) cycles sine/cosine, respectively. This filtering included a linear trend removal. Finally, functional runs were co-registered and aligned to the anatomical scan using an affine transformation (9 parameters) and z-scored to eliminate signal offsets and inter-run variance.

REGION-OF-INTEREST DEFINITION Different regions-of-interest (ROIs) were defined in each subject. Using the eccentricity and polar angle maps that resulted from the pRF fitting and were projected onto the inflated surface reconstructions, we defined region-of-interests (ROIs) for bilateral visual areas V_1 , V_2 , and V_3 . The resulting surface patches from the left and right hemisphere were projected back into volume space (from -1 mm to +3 mm from the segmented grey/white matter boundary). Volume ROIs were then defined for V_1 , V_2 , V_3 , and a combined ROI ($V_1V_2V_3$). With data from the VWFA localizer we defined a VWFA volume ROI ($q(\text{FDR}) < .05$) in the left hemisphere of each subject contrasting against neighbouring regions like fusiform face area (Stigliani, Weiner, & Grill-Spector, 2015).

MULTI VOXEL PATTERN ANALYSIS We computed a multi-voxel pattern analysis (MVPA) for each ROI. Similar to a previous study (Emmerling et al., 2016), functional data from the experimental imagery runs were split into training and testing datasets in a leave-one-run-out procedure that allowed to cross-validate the performance of the classifier. In each ROI and cross-validation, using data from the training dataset, the 1000 voxels with the highest F-values (omnibus test) were pre-selected as features to reduce the high number of potential features (voxels) in each ROI that were available due to the high spatial resolution. Using data from these pre-selected voxels, we extracted an average of 6 seconds of data (2 TRs; time window from 2 TRs until 4 TRs after trial onset) as final features. In a one-versus-one 4-class classification (winner is defined by the maximum number of votes in all binary classifications) employing a support vector machine (SVM; LIBSVM implementation in PyMVPA; Chang and Lin, 2011), the imagined letters in each trial of the left-out test dataset were decoded. To test the

probability of the classification accuracies we repeated the whole analysis 1000 times with scrambled labels and obtained a distribution of classification accuracies under the null hypothesis. Real classification accuracies were tested against this null distribution with an exact permutation test resulting in p-values for each subject and ROI.

VISUAL RECONSTRUCTION OF IMAGINED AND PERCEIVED LETTERS For each subject, we reconstructed the visual field for each TR during the perception runs and during the experimental imagery runs into an estimated image of the visual field (VFI). Taking into account the activation in all voxels within one ROI that had an R value above the 99th percentile we used the x , y , and sigma parameters that were fitted to each voxel during the pRF mapping to project their activation to a 150×150 pixel image using a Gaussian kernel. To account for overrepresentations of certain areas in the visual field we divided the resulting activations by the number of voxels that influenced a pixel. To obtain an image with a higher signal-to-noise ratio we averaged the VFI for every letter using data of two TRs starting at one, two, or three TRs after trial onset (three shifts). This time-window analysis was performed to investigate the temporal characteristics of the visual activations during imagery and perception. For the *Imagery* VFI we averaged data from all 32 imagery trials per letter (four runs), for the *Perception* VFI we averaged data from all eight perception trials per letter (one run). Additionally, we assembled an *Idealised* dataset ('Ide') by importing screenshots of the presented letters cropped and transformed to the dimensions of the 150×150 pixel image.

CORRELATION OF AVERAGE VISUAL FIELD IMAGES For every subject, ROI, and shift we computed a dissimilarity matrix of the

vectorised VFIs for every letter. Dissimilarity matrices were also computed between the average images of the *Perception* and *Imagery* ('Per vs. Ima'), *Imagery* and *Idealised* ('Ima vs. Ide'), and *Idealised* and *Perception* ('Per vs. Ide') datasets.

DECODING OF SINGLE IMAGERY TRIALS By comparing the VFIs of single trials during the experimental imagery runs with the average VFIs of the *Idealised*, *Perception*, and *Imagery* dataset we decoded the imagined letter for each trial. The vectorised VFI of the single trial (average of two TRs, see above) was correlated with all four average letter VFIs of the respective dataset. Subsequently, the obtained correlation vector was compared to the dissimilarity matrix vector (row) obtained in the previous steps by means of a second-level correlation to account for inherent similarities between the letter representations. The highest second-level correlation was chosen as the classified letter. For the *Imagery* dataset we employed a leave-one-run-out cross-validation procedure to generate template average VFIs from three of the four imagery runs and compare them to the single-trial VFIs from the left-out run. Results for the *Imagery* dataset show the averages of these cross-validations.

DECODING FROM VISUAL WORD FORM AREA We analysed z-scored raw voxel data from the defined VWFA ROI to decode single imagery trials in a similar procedure as for the VFIs. The multi-voxel patterns within the ROI were averaged and formed *Perception* and *Imagery* dataset templates. After computing dissimilarity matrices for both datasets (see above) they were used to decode the multi-voxel pattern in single letter imagery trials. For the *Imagery* dataset, a cross-validation procedure was performed as outlined before. In addition to a classification with all four letters (4-class)

we computed a pooled decoding analysis (2-class) for the letters 'H' and 'T' versus 'S' and 'C'. When combining letters with similar features (edges versus curves), feature-selective activations in VWFA could enhance the performance of this analysis.

PERMUTATION TESTING We tested the results from the pipeline for VFIs and the VWFA voxel data (outlined above) for significance by means of permutation testing. The whole procedure was repeated 1000 times with randomly scrambled labels (between trials; within runs). The obtained distribution of results for each analysis was compared to the empirical results using an exact permutation test yielding p-values for each subject, ROI, shift, dataset, and analysis, respectively.

4.3 RESULTS

BEHAVIOURAL DATA VVIQ scores and OSIVQ scores (all three scales) are shown in [Figure 4.2](#) for all subjects. Subjects reported that they tried to maintain the afterimage of the fading out stimulus as a strategy to enforce vivid and accurate letter imagery. Average accuracies of the responses recorded during the probing phases improved over the three training sessions for both, training runs with the visible (but fading out) letters as well as for the training runs without visible letters (see [Figure 4.3](#)).

MULTI VOXEL PATTERN ANALYSES Results from the MVPAs are shown in [Figure 4.4](#). Classification accuracies in both subjects and ROIs did not reach significance as assessed by exact permutation tests.

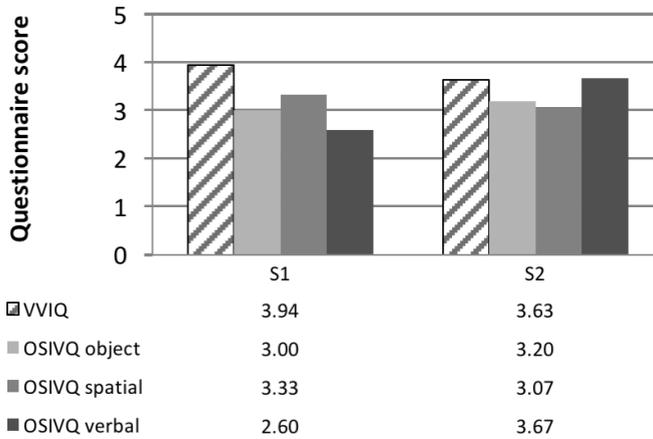


Figure 4.2 Vividness of Visual Imagery Questionnaire (VVIQ) and Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) scores. VVIQ scores and OSIVQ scores (with the subscales for “object”, “spatial”, and “verbal” imagery styles) are shown for both subjects (S₁, S₂).

VISUAL RECONSTRUCTION OF IMAGINED AND PERCEIVED LETTERS Reconstructed VFIs for each subject, ROI, and letter are shown for imagery and perception in [Figure 4.5](#) (shift +2 is visualised; for other shifts, see supplementary material). In both subjects, perception VFIs bear strong resemblance with the actual presented letter shape (see bottom row of [Figure 4.5](#); *Idealised* dataset). Especially, in the ROIs ‘V₁’ and ‘V₁V₂V₃’ there are strong similarities between the VFIs for imagery and perception. While not as clear as the perceived reconstructed VFIs, reconstructed imagines letter shapes differ clearly from each other.

LETTER VISUAL FIELD IMAGE DISSIMILARITY MATRICES The correlations between all letters for the three datasets *Imagery*, *Perception*, and *Idealised*, and combinations thereof, are shown for ‘shift +2’ in [Figure 4.6](#) (subject S₁) and [Figure 4.7](#) (subject S₂). In

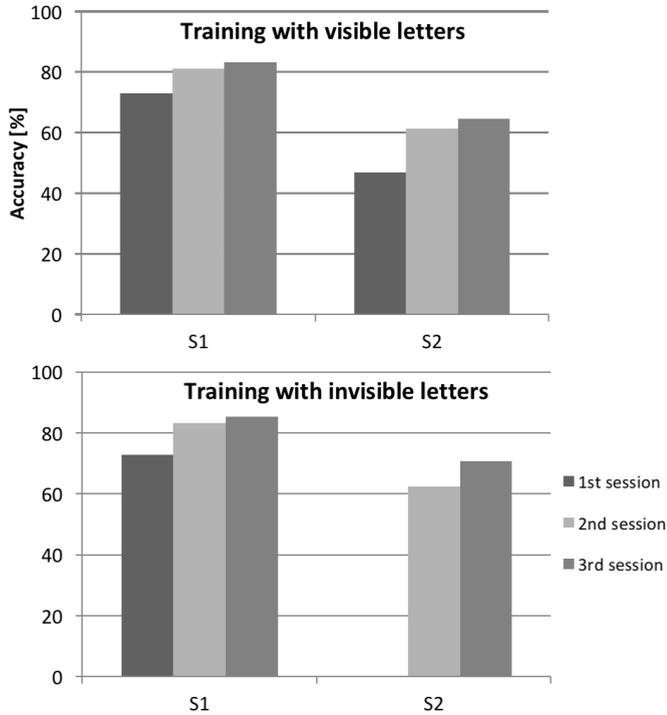


Figure 4.3 Average accuracies for the probing phases during training. Averaged accuracies for the probing phases in the training task are shown for both subjects (S1, S2) over the three training sessions. Data from the two runs of training with visible (but fading out) letters (left) and data from the two runs of training without visible letters are shown, respectively. Note: subject S2 did not perform the training with invisible letters in the first training session.

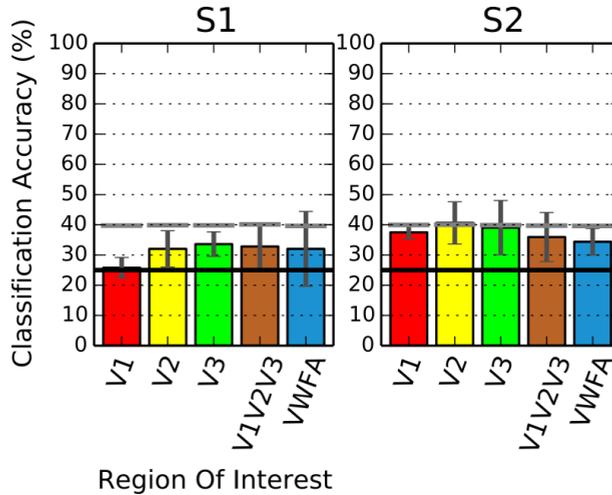


Figure 4.4 Multi-voxel pattern analysis classification accuracies for five regions-of-interest. Classification accuracies of subjects S1 and S2 are reported for the ROIs V1, V2, V3, V1V2V3, and VWFA in time window shifts '+2'. Black horizontal lines indicate the mean classification accuracies across all 1000 permutations; grey horizontal lines demarcate the 95th percentile of permutation classification accuracies. Note: Colours are consistent across Figures.

all dissimilarity matrices of the dataset *Imagery* ('Ima vs. Ima') eight correlations were significant. For the dataset *Perception* ('Per vs. Per') two correlations were significant. When comparing the average letter VFIs of *Perception* and *Imagery* ('Per vs. Ima') 31 correlations became significant, in the comparison of *Imagery* and *Idealised* ('Ima vs. Ide') 34 correlations became significant, and in the comparison of *Perception* and *Idealised* ('Per vs. Ide') 51 correlations became significant (all $p < .05$; as tested against 1000 permutations).

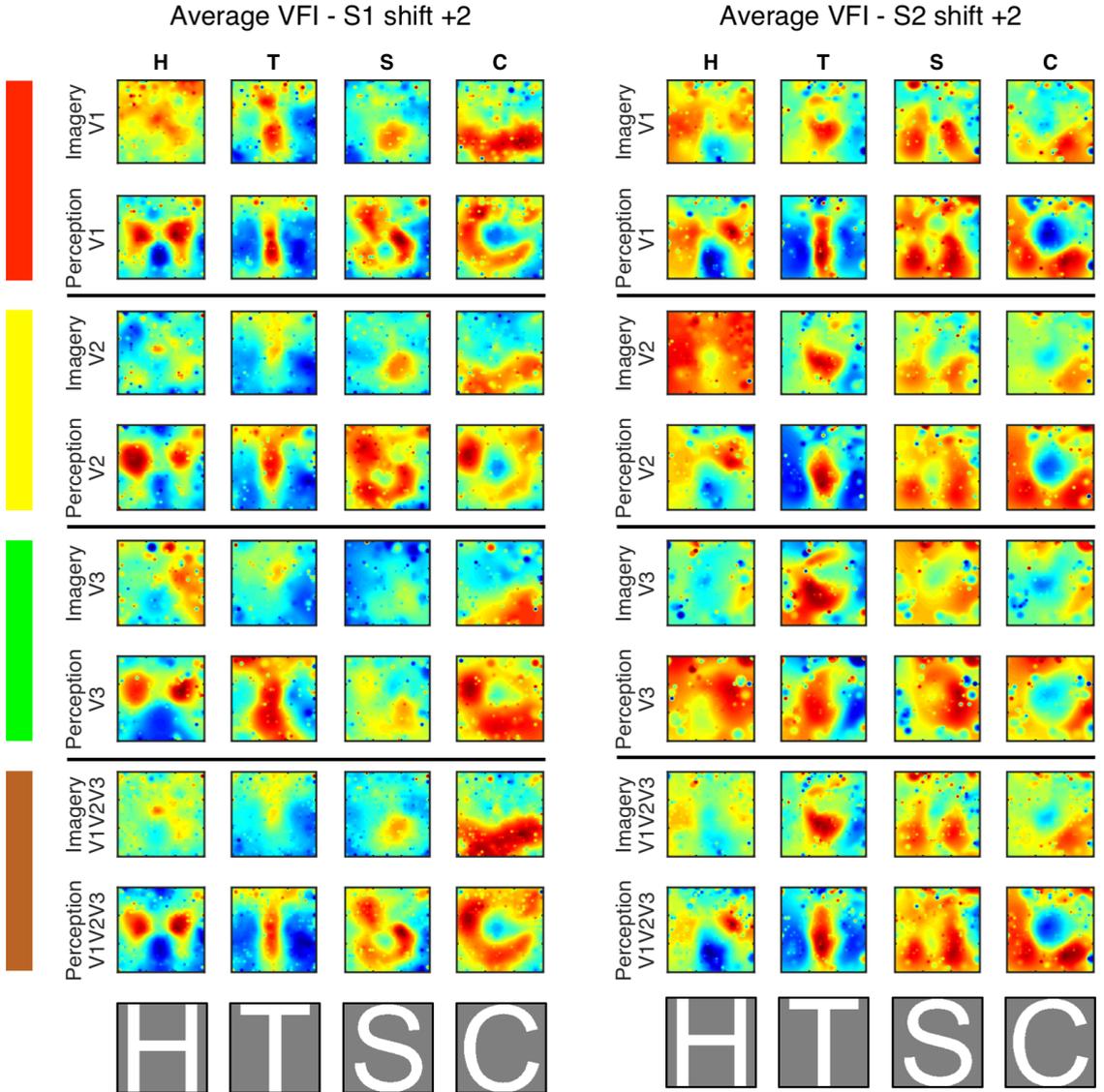


Figure 4.5 Reconstructed averaged visual field images (VFIs). Reconstructed VFIs are visualised for subjects S1 (left) and S2 (right) using data from time window 'shift +2' (from +2 TRs until +4 TRs after trial onset). VFIs for every letter (H, T, S, and C) are constructed for each ROI (V1, V2, V3, and V1V2V3; colours are consistent across Figures) and for both, imagery (top rows) and perception (bottom rows). In the bottom row the presented letter shapes are shown for comparison ('idealised' dataset). Note: Averages for imagery VFIs involved data of 32 trials per letter, averages for perception VFIs involved data of 8 trials per letter. Colours in each plot are scaled to the respective range of values to allow for better visual inspection.

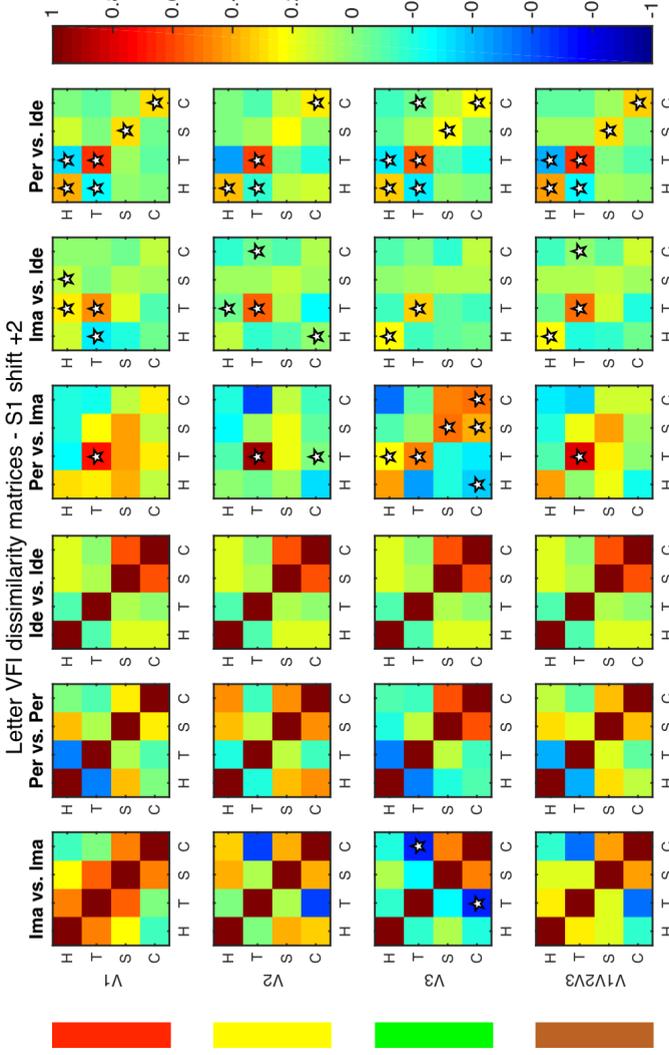


Figure 4.6 Letter visual field image (VFI) dissimilarity matrices for different datasets and dataset combinations in subject S1. Dissimilarity matrices are plotted for subject S1 using data from time window ‘shift +2’ (from +2 TRs until +4 TRs after trial onset) in each ROI (V_1 , V_2 , V_3 , and V_{1V2V3} ; colours are consistent across Figures). Dissimilarity matrices between all four imagined or perceived letters (H, T, S, and C) are shown for the three datasets *Imagery* (‘Ima vs. Ima’; first column), *Perception* (‘Per vs. Per’; second column), and *Idealised* (‘Ide vs. Ide’; third column). Additionally, dissimilarity matrices between all letters are shown for cross-dataset combinations *Perception* versus *Imagery* (‘Per vs. Ima’; fourth column), *Imagery* versus *Idealised* (‘Ima vs. Ide’; fifth column), and *Perception* versus *Idealised* (‘Per vs. Ide’; sixth column). Colours correspond to Pearson correlation indices (see colour bar). Stars indicate significant correlation ($p < .05$) as assessed by testing against 1000 permutations. Note: Single dataset dissimilarity matrices (first three columns ‘Ima vs. Ima’, ‘Per vs. Per’, ‘Ide vs. Ide’) are, by definition, symmetric. The dissimilarity matrix for the *Idealised* dataset (‘Ide vs. Ide’, third column) is plotted for every ROI (all rows) to allow for easier visual comparison.

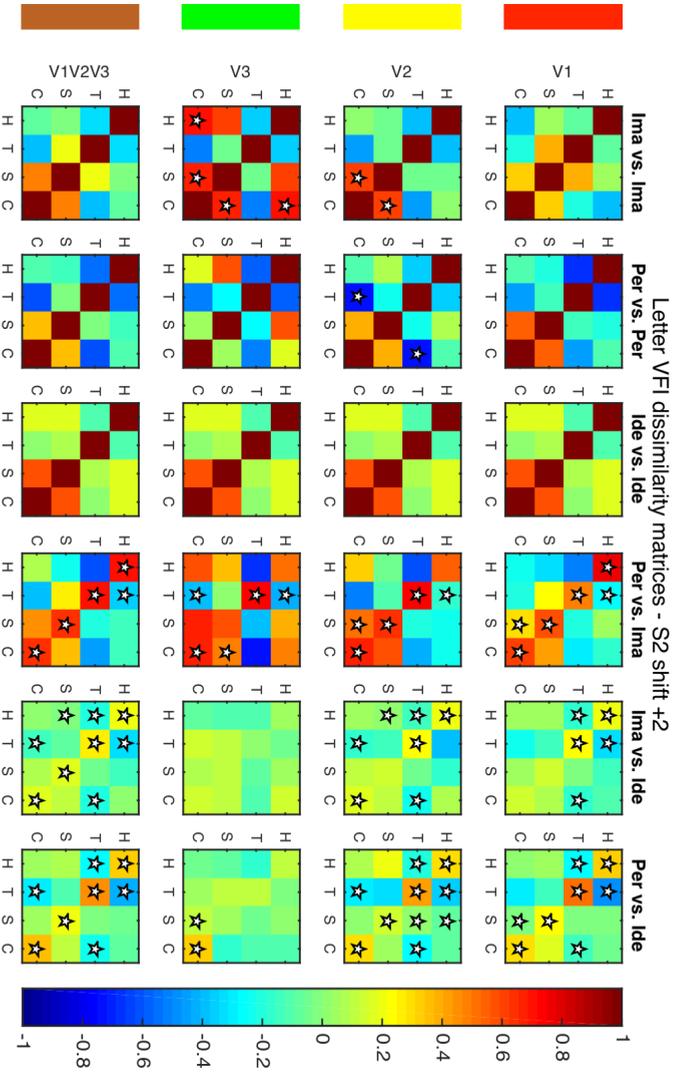


Figure 4.7 Letter visual field image (VFI) dissimilarity matrices for different datasets and dataset combinations in subject S2. For a detailed description, please see caption of Figure 4.6.

DECODING OF SINGLE IMAGERY TRIALS Decoding accuracies for all ROIs, shifts, and template datasets (*Idealised*, *Perception*, and *Imagery*) are shown in [Figure 4.8](#) (subject S1) and [Figure 4.9](#) (subject S2). When using the *Idealised* average letter VFIs and their respective self-correlation vectors as a template, 22 decoding accuracies were significant. Twenty-three decoding accuracies for the *Perception* templates and no decoding accuracy for the *Imagery* template were significant (all $p < .05$; as tested against 1000 permutations).

DECODING OF SINGLE IMAGERY TRIALS IN VWFA When using raw (z-scored voxel) data from the VWFA ROI to decode single imagery trials, two decoding accuracies were significant for the *Perception* dataset (shift '+2' for subjects S01 and S02) when pooling similar letters ('H' and 'T' versus 'S' and 'C'). No classification accuracies were significant for the *Imagery* dataset (all $p < .05$; as tested against 1000 permutations). Results for the single-trial decoding in VWFA are shown in [Figure 4.10](#).

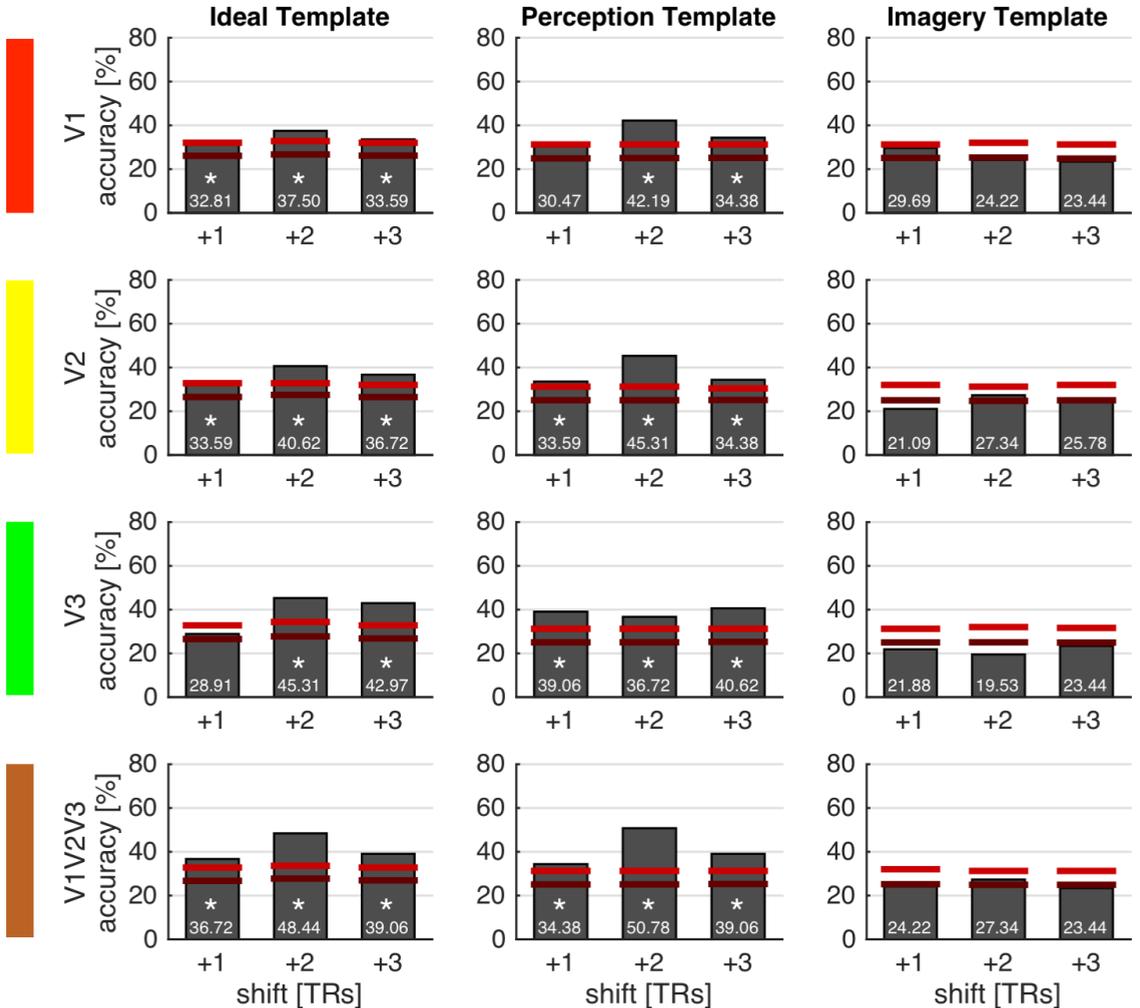


Figure 4.8 Single-trial letter visual field image (VFI) decoding results for subject S1. Classification accuracies are reported for the ROIs V_1 , V_2 , V_3 , and $V_1V_2V_3$ and the three time window shifts ('+1', '+2', '+3') when using the average letter VFIs of the *Idealised* (first column), *Perception* (second column), or *Imagery* (third column) dataset as templates (taking into account their respective self-correlation). Classification accuracies are reported at the bottom of each bar and asterisks indicate significant classification accuracies ($p < .05$) as assessed by testing against 1000 permutations. Dark red horizontal lines indicate the mean classification accuracies across all permutations; bright red horizontal lines demarcate the 95th percentile of permutation classification accuracies. Note: Colours are consistent across Figures.

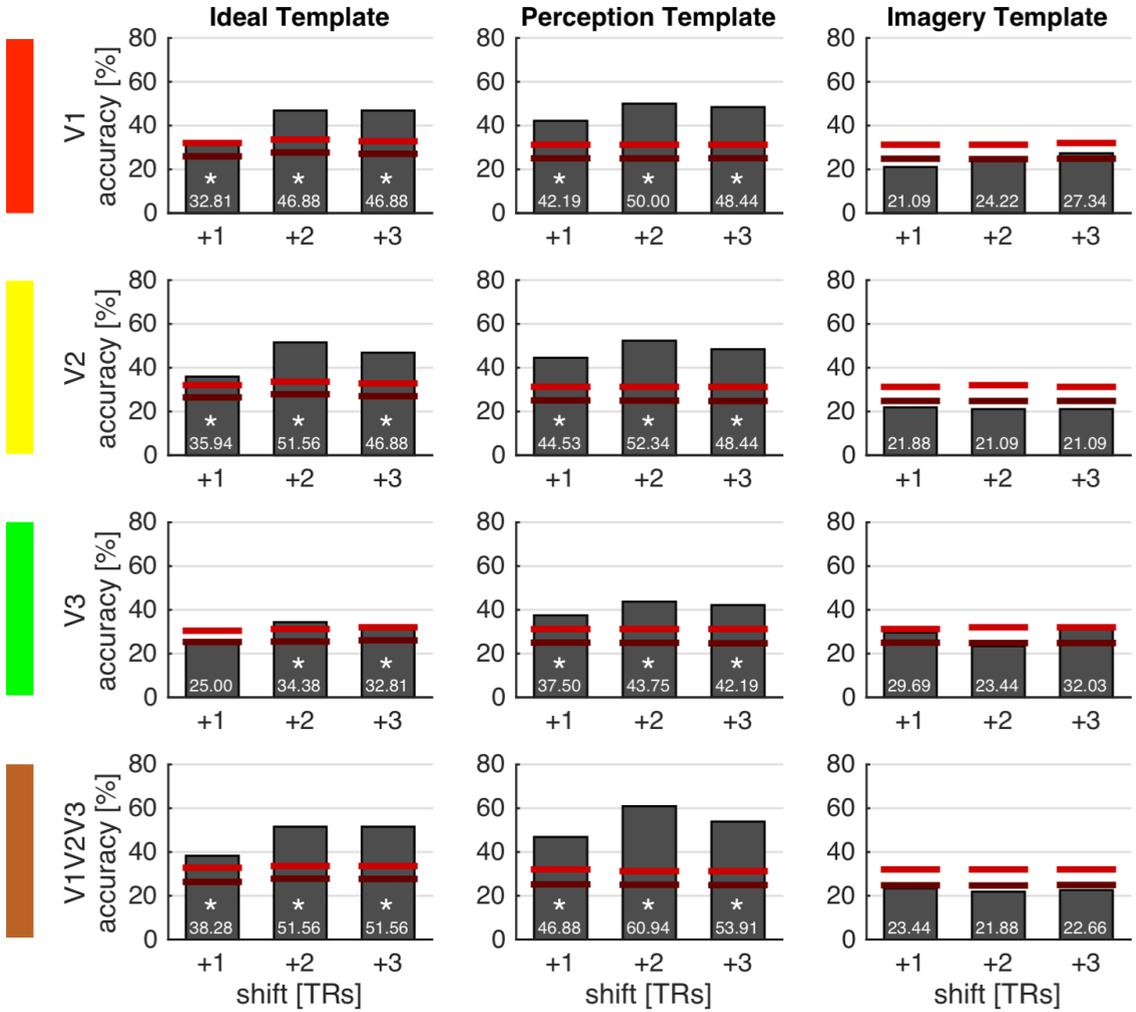
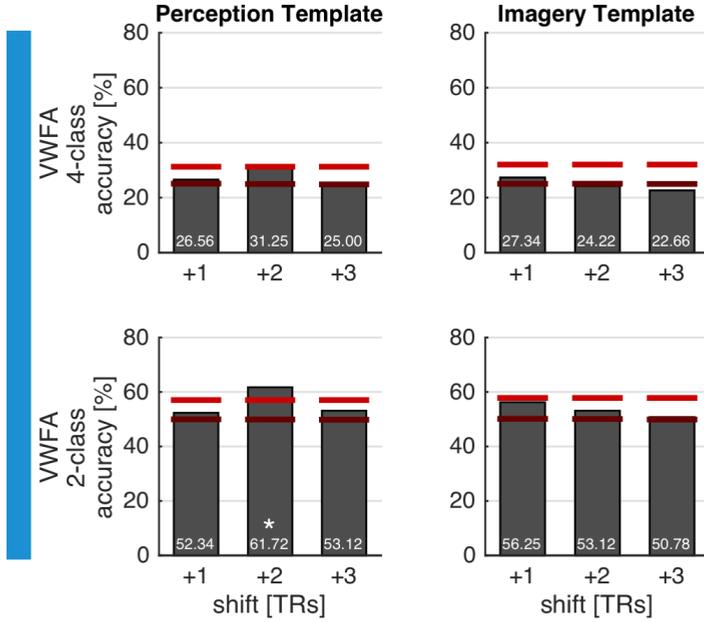
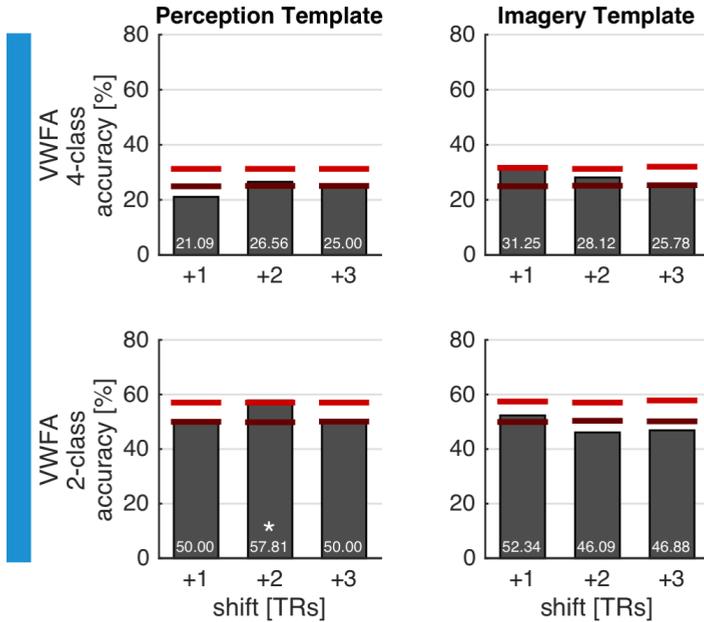


Figure 4.9 Single-trial letter visual field image (VFI) decoding results for subject S2. For a detailed description, please see caption of Figure 4.8.



(a) Single-trial letter decoding in VWFA for subject S1



(b) Single-trial letter decoding in VWFA for subject S2

Figure 4.10 Results from single-trial letter decoding in visual word form area (page 140). Classification accuracies are reported for (a) subject S1 and (b) subject S2 regarding the three time window shifts ('+1', '+2', '+3') when using the z-scored voxel data from the visual word form area (VWFA) region-of-interest (ROI). Classifications are performed for 4-class (first rows in (a) & (b); all letters) and 2-class (second rows in (a) & (b); 'H' and 'T' versus 'S' and 'C') decoding. Average multi-voxel patterns for each letter of the *Perception* (first column), or *Imagery* (second column) dataset are used as templates (taking into account their respective self-correlation). Classification accuracies are reported at the bottom of each bar and asterisks indicate significant classification accuracies ($p < .05$) as assessed by testing against 1000 permutations. Dark red horizontal lines indicate the mean classification accuracies across all permutations; bright red horizontal lines demarcate the 95th percentile of permutation classification accuracies. Note: Colour are consistent across Figures.

4.4 DISCUSSION

In this study, we reconstructed and decoded imagined letter shapes from submillimetre resolution fMRI data. Using a computed pRF mapping for individual subjects, we reconstructed the visual field during perception and imagery of four different letter shapes and found a remarkable similarity to the presented letter shapes, especially in the earliest visual areas. We were able to significantly decode reconstructed visual field images (VFIs) based on single imagery trials using perception VFIs or even ideal letter shapes (presented stimuli) as reference. Furthermore, we found significant overlap between reconstructed VFIs of perception data, imagery data, and the actually presented stimuli. Finally, we successfully classified between morphologically edgy ('H' and 'T') and morphologically curvy ('S' and 'C') letters using multi-voxel pattern data from VWFA.

We show, for the first time, reconstructed VFIs that carry recognisable content of a complex mental image, even on a single trial. Previous research has demonstrated functional retinotopic activation of early visual areas during visual mental imagery (Albers et al., 2013; Klein et al., 2004; Naselaris et al., 2014; Slotnick et al., 2005). Stokes et al. (2009) were able to decode whether subjects imagined the letter 'X' or the letter 'O' from fMRI data of visual cortex using MVPA. Their results demonstrated that it is possible to decode mental imagery of letters, while leaving unclear exactly what a mental image of an imagined letter would look like. By decoding the reconstructed VFIs from single imagery trials we put a hard test to the hypothesis that visual mental imagery and perception share the same neural correlates: We computed pRF mappings of visual stimuli and reconstructed imagined stimuli using these (perception-based) parameters. In addition, we classified single tri-

als using reference images from the *Perception* dataset or even the *Idealised* dataset. It has to be stressed that classifying reconstructed VFIs by using correlations is a much more conservative test than the usage of MVPA on fMRI data directly. Interestingly, that single-trial decoding was not successful when we used averaged imagery VFIs as reference images. This was probably due to high inter-trial fluctuations of imagery performance that led to a comparably low signal-to-noise ratio of the *Imagery* reference dataset.

The reconstruction and decoding of visual mental imagery not only sheds light on the neural correlates of mental imagery and the interaction with perceptual processes, but is also of interest for the development of brain-computer interfaces (BCIs). In a BCI, the decoded type or content of the imagery is used as an input for a computer program to express the BCI user's intent (Birbaumer & Cohen, 2007). When other means of interaction or communication with the environment have become impossible due to certain diseases like amyotrophic lateral sclerosis (ALS), BCI systems can offer great help to patients. BCI communication systems have mostly focused on creating some kind of alphabetical speller that allows for writing words and sentences using mental imagery (Birbaumer et al., 1999; Sorger, Reithler, Dahmen, & Goebel, 2012). Current BCI speller systems based on internal cognitive processes, however, do not offer a close connection between imagery content and intended letter in that users of these systems have to perform arbitrary mental imagery tasks (e. g., mental spatial navigation or inner speech) that allow for good classification accuracies. Although our results only include data from two healthy subjects, this study provides first evidence that a more natural and content-based BCI speller system is feasible, decoding imagined letters from visual cortex rather than having the BCI user encode the intended letter with arbitrary imagery tasks. A more direct relationship between im-

*Brain-
computer
interfaces*

agery content and spelled letter would be desirable, especially for novice BCI speller users. Moreover, our successful single-trial imagery decoding using screenshots of the presented letter shapes (*Idealised* dataset) shows, that it would be possible to decode imagined content in a new subject with only recording a standard pRF mapping. This would eliminate the need for prior training periods necessary for most BCI systems (Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002).

Outlook This study is a proof of concept, demonstrating the feasibility of reconstructing imagined letters from high-resolution neuroimaging data of the visual cortex. In the future, reconstruction and single-trial decoding could be improved by using more sophisticated reconstruction methods (Yargholi & Hossein-Zadeh, 2016) or image recognition algorithms (Hinton, Osindero, & Teh, 2006) to classify the VFIs. This would potentially increase single-trial decoding accuracies as well as the number of letters that can be distinguished.

REFERENCES

- Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013, August). Shared representations for working memory and mental imagery in early visual cortex. *Current biology : CB*, 23(15), 1427–31. doi:[10.1016/j.cub.2013.05.065](https://doi.org/10.1016/j.cub.2013.05.065)
- Andersson, J. L. R., Skare, S., & Ashburner, J. (2003, October). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage*, 20(2), 870–88. doi:[10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7)
- Bartolomeo, P. (2002). The relationship between visual perception and visual mental imagery: a reappraisal of the neuropsychological evidence. *Cortex*, 357–378. doi:[10.1016/S0010-9452\(08\)70665-8](https://doi.org/10.1016/S0010-9452(08)70665-8)
- Birbaumer, N. & Cohen, L. G. (2007, March). Brain-computer interfaces: communication and restoration of movement in paralysis. *The Journal of physiology*, 579(Pt 3), 621–36. doi:[10.1113/jphysiol.2006.125633](https://doi.org/10.1113/jphysiol.2006.125633)
- Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., ... Flor, H. (1999, March). A spelling device for the paralysed. *Nature*, 398(6725), 297–8. doi:[10.1038/18581](https://doi.org/10.1038/18581)
- Blazhenkova, O. & Kozhevnikov, M. (2009, July). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, 23(5), 638–663. doi:[10.1002/acp.1473](https://doi.org/10.1002/acp.1473)
- Chang, C.-C. & Lin, C.-J. (2011, April). Libsvm. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–27. doi:[10.1145/1961189.1961199](https://doi.org/10.1145/1961189.1961199)
- Cichy, R. M., Heinzle, J., & Haynes, J.-D. (2012, February). Imagery and perception share cortical representations of content and location. *Cerebral cortex (New York, N.Y. : 1991)*, 22(2), 372–80. doi:[10.1093/cercor/bhr106](https://doi.org/10.1093/cercor/bhr106)
- Dumoulin, S. O. & Wandell, B. A. (2008, January). Population receptive field estimates in human visual cortex. *NeuroImage*, 39(2), 647–60. doi:[10.1016/j.neuroimage.2007.09.034](https://doi.org/10.1016/j.neuroimage.2007.09.034)
- Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A., & Goebel, R. (2016, January). Decoding the direction of imagined visual motion using 7T ultra-high field fMRI. *NeuroImage*, 125, 61–73. doi:[10.1016/j.neuroimage.2015.10.022](https://doi.org/10.1016/j.neuroimage.2015.10.022)
- Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004, July). Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Brain*

- research. *Cognitive brain research*, 20(2), 226–41. doi:[10.1016/j.cogbrainres.2004.02.012](https://doi.org/10.1016/j.cogbrainres.2004.02.012)
- Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of Functional Image Analysis Contest (FIAC) data with BrainVoyager QX: From single-subject to cortically aligned group General Linear Model analysis and self-organizing group Independent Component Analysis. *Human Brain Mapping*, 27(5), 392–401. doi:[10.1002/hbm.20249](https://doi.org/10.1002/hbm.20249)
- Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998, May). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *The European journal of neuroscience*, 10(5), 1563–73. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9751129>
- Harrison, S. A. & Tong, F. (2009, April). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–5. doi:[10.1038/nature07832](https://doi.org/10.1038/nature07832)
- Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006, July). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527–54. doi:[10.1162/neco.2006.18.7.1527](https://doi.org/10.1162/neco.2006.18.7.1527)
- Ishai, A., Ungerleider, L. G., & Haxby, J. V. (2000). Distributed neural systems for the generation of visual images. *Neuron*, 28(3), 979–90. doi:[10.1016/S0896-6273\(00\)00168-9](https://doi.org/10.1016/S0896-6273(00)00168-9)
- Johnson, M. R. & Johnson, M. K. (2014, January). Decoding individual natural scene representations during perception and imagery. *Frontiers in human neuroscience*, 8(February), 59. doi:[10.3389/fnhum.2014.00059](https://doi.org/10.3389/fnhum.2014.00059)
- Klein, I., Dubois, J., Mangin, J. F., Kherif, F., Flandin, G., Poline, J.-B., ... Le Bihan, D. (2004). Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, 22(1), 26–31. doi:[10.1016/j.cogbrainres.2004.07.006](https://doi.org/10.1016/j.cogbrainres.2004.07.006)
- Kosslyn, S. M. & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological bulletin*, 129(5), 723–746. doi:[10.1037/0033-2909.129.5.723](https://doi.org/10.1037/0033-2909.129.5.723)
- Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997, November). Neural systems shared by visual imagery and visual perception: a positron emission tomography study. *NeuroImage*, 6(4), 320–34. doi:[10.1006/nimg.1997.0295](https://doi.org/10.1006/nimg.1997.0295)
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006, March). *The Case for Mental Imagery*. Oxford University Press. doi:[10.1093/acprof:oso/9780195179088.001.0001](https://doi.org/10.1093/acprof:oso/9780195179088.001.0001)

- Kosslyn, S. M., Thompson, W. L., Wraga, M., & Alpert, N. M. (2001, August). Imagining rotation by endogenous versus exogenous forces: distinct neural mechanisms. *Neuroreport*, 12(11), 2519–25. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11496141>
- Kriegeskorte, N. & Goebel, R. (2001). An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical mr volumes. *NeuroImage*, 14(2), 329–346. doi:10.1006/nimg.2001.0831
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2012, February). Disentangling visual imagery and perception of real-world objects. *NeuroImage*, 59(4), 4064–73. doi:10.1016/j.neuroimage.2011.10.055
- Marks, D. F. (1973, February). Visual imagery differences in the recall of pictures. *British journal of psychology (London, England : 1953)*, 64(1), 17–24. doi:10.1111/j.2044-8295.1973.tb01322.x
- Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.-F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage*, 49(2), 1271–1281. doi:10.1016/j.neuroimage.2009.10.002
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003, July). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7(7), 293–299. doi:10.1016/S1364-6613(03)00134-7
- Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004, November). Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cerebral cortex (New York, N.Y. : 1991)*, 14(11), 1256–65. doi:10.1093/cercor/bhh087
- Miyawaki, Y., Uchida, H., Yamashita, O., Sato, M.-a., Morito, Y., Tanabe, H. C., . . . Kamitani, Y. (2008, December). Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron*, 60(5), 915–29. doi:10.1016/j.neuron.2008.11.004
- Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E. J., Strupp, J. P., Harel, N., & Uğurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain FMRI. *Magnetic Resonance in Medicine*, 63(5), 1144–1153. doi:10.1002/mrm.22361
- Naselaris, T. & Lynam, K. (2014). Objective mental images. In *Poster presented at neuroscience 2014*.
- Naselaris, T., Olman, C. A., Stansbury, D. E., Uğurbil, K., & Gallant, J. L. (2014, October). A voxel-wise encoding model for early visual areas decodes

- mental images of remembered scenes. *NeuroImage*, 105, 215–228. doi:[10.1016/j.neuroimage.2014.10.018](https://doi.org/10.1016/j.neuroimage.2014.10.018)
- O’Craven, K. M. & Kanwisher, N. (2000, November). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of cognitive neuroscience*, 12(6), 1013–23. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11177421>
- Pearson, J., Naselaris, T., Holmes, E. A., & Kosslyn, S. M. (2015, October). Mental Imagery: Functional Mechanisms and Clinical Applications. *Trends in Cognitive Sciences*, 19(10), 590–602. doi:[10.1016/j.tics.2015.08.003](https://doi.org/10.1016/j.tics.2015.08.003)
- Peirce, J. W. (2007, May). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2), 8–13. doi:[10.1016/j.jneumeth.2006.11.017](https://doi.org/10.1016/j.jneumeth.2006.11.017)
- Reddy, L., Tsuchiya, N., & Serre, T. (2010, April). Reading the mind’s eye: decoding category information during mental imagery. *NeuroImage*, 50(2), 818–25. doi:[10.1016/j.neuroimage.2009.11.084](https://doi.org/10.1016/j.neuroimage.2009.11.084)
- Schoenmakers, S., Barth, M., Heskes, T., & van Gerven, M. a. J. (2013, December). Linear reconstruction of perceived images from human brain activity. *NeuroImage*, 83, 951–61. doi:[10.1016/j.neuroimage.2013.07.043](https://doi.org/10.1016/j.neuroimage.2013.07.043)
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., ... Tootell, R. B. H. (1995, May). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science (New York, N.Y.)* 268(5212), 889–93. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7754376>
- Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2005, October). Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral cortex (New York, N.Y. : 1991)*, 15(10), 1570–83. doi:[10.1093/cercor/bhio35](https://doi.org/10.1093/cercor/bhio35)
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., ... Matthews, P. M. (2004, January). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1, S208–19. doi:[10.1016/j.neuroimage.2004.07.051](https://doi.org/10.1016/j.neuroimage.2004.07.051)
- Sorger, B., Reithler, J., Dahmen, B., & Goebel, R. (2012, July). A real-time fMRI-based spelling device immediately enabling robust motor-independent communication. *Current biology : CB*, 22(14), 1333–8. doi:[10.1016/j.cub.2012.05.022](https://doi.org/10.1016/j.cub.2012.05.022)

- Stigliani, A., Weiner, K. S., & Grill-Spector, K. (2015). Temporal Processing Capacity in High-Level Visual Cortex Is Domain Specific. *Journal of Neuroscience*, 35(36), 12412–12424. doi:[10.1523/JNEUROSCI.4822-14.2015](https://doi.org/10.1523/JNEUROSCI.4822-14.2015)
- Stokes, M., Saraiva, A., Rohenkohl, G., & Nobre, A. C. (2011, June). Imagery for shapes activates position-invariant representations in human visual cortex. *NeuroImage*, 56(3), 1540–5. doi:[10.1016/j.neuroimage.2011.02.071](https://doi.org/10.1016/j.neuroimage.2011.02.071)
- Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009, February). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(5), 1565–72. doi:[10.1523/JNEUROSCI.4657-08.2009](https://doi.org/10.1523/JNEUROSCI.4657-08.2009)
- Thirion, B., Duchesnay, E., Hubbard, E., Dubois, J., Poline, J.-B., Lebihan, D., & Dehaene, S. (2006, December). Inverse retinotopy: inferring the visual content of images from brain activation patterns. *NeuroImage*, 33(4), 1104–16. doi:[10.1016/j.neuroimage.2006.06.062](https://doi.org/10.1016/j.neuroimage.2006.06.062)
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002, June). Brain-computer interfaces for communication and control. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 113(6), 767–91. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12048038>
- Yargholi, E. & Hossein-Zadeh, G. A. (2016). Reconstruction of digit images from human brain fMRI activity through connectivity informed Bayesian networks. *Journal of Neuroscience Methods*, 257, 159–167. doi:[10.1016/j.jneumeth.2015.09.032](https://doi.org/10.1016/j.jneumeth.2015.09.032)

Chapter 5

General discussion

This thesis investigated the functional architecture of different cortical areas in the human visual system in detail using ultra-high resolution functional magnetic resonance imaging (fMRI). For perception and imagery, specific fine-grained functional organisations within different visual areas were revealed. [Chapter 2](#) focussed on human middle temporal area (hMT). For the first time, a dual tuning – for binocular disparity and for direction-of-motion – could be mapped in an extrastriate area. These results were not only the first to show tuning for binocular disparity in hMT, but more importantly, demonstrated how cortical areas in the human visual system are simultaneously tuned for several features. [Chapter 3](#), equally investigating the functional architecture of hMT, took the preceding results one step further in focusing not only on visual perception, but also imagery. The results regarding the successful classification of different imagined directions-of-motion revealed that sub-categorical decoding of complex mental imagery is possible even in single trials. These results not merely answer the question of *whether* imagery effects can be shown in early visual cortex but answer the question of *what* was imagined. The levels of accuracy we achieved and the extent of brain activations in single subjects we demonstrated can become the building blocks for future brain-computer interface (BCI) that decode the intent of the BCI user much more directly. Finally, in [Chapter 4](#) the retinotopic organisation of visual areas was leveraged when imagined letter shapes were reconstructed from fMRI data of early visual areas and, thereafter, successfully classified. The results in [Chapter 4](#) are the first to show a clear reconstruction of a complex mental image. We observed strong similarities between brain activations in visual cortex when subjects perceived a letter shape versus when they imagined it. These similarities allowed the rather conservative test of classifying the reconstructed visual field images during

imagery using data from perception runs. The findings presented in this thesis add to three major fields of visual neuroscience as discussed in the next paragraphs.

*Feature-level
functional
organisation
of the visual
system*

All studies presented in this thesis reveal feature representations within visual areas. Investigating the feature-level functional organisation of the human brain is relatively novel to cognitive neuroscience as neuroimaging methods did not provide sufficient spatial resolution to probe these representations until recently. By using ultra-high field fMRI, yielding submillimetre resolution functional data, we could ‘zoom’ into the fine-grained functional organisation of visual areas. We were able to go beyond a mere localisation of functional brain areas towards mapping their internal functional architecture. With the demonstration of two distinct feature tunings in hMT ([Chapter 2](#)) we replicated findings from macaque cortex (DeAngelis & Newsome, 1999) and found for the first time a dual feature tuning in an extrastriate area of human visual cortex. This second finding in particular surpasses previous studies that mapped tunings to one visual feature (Goncalves et al., 2015; Yacoub, Harel, & Ugurbil, 2008; Zimmermann et al., 2011) or report sensitivity to multiple feature dimensions (Nasr, Polimeni, & Tootell, 2016) without mapping the precise feature organisation. Our results are especially important, as they will help to finally unravel the complex multi-purpose computational pipeline of the visual system that is needed to accomplish the tasks the visual system is confronted with on a daily basis.

*Decoding and
reconstruct-
ing mental
images*

As it becomes feasible to conduct neuroimaging experiments targeting top-down influences on visual areas – for instance by employing layer-specific sampling (Kok, Bains, van Mourik, Norris, & de Lange, 2016; Muckli et al., 2015) – they slowly attract more and more attention in visual neuroscience (Gilbert & Sigman, 2007). With the results presented in [Chapter 3](#) and [Chapter 4](#),

this thesis contributes to the understanding of the processes underlying one of the most prominent forms of top-down influences in the visual system: mental imagery. With decoding not only the type of visual mental imagery, but also its actual content, we begin to understand its functional underpinnings. This is an important milestone as it allows cognitive neuroscience to tap into the mental representations that humans form during cognitive processes. Previously, studies on mental imagery were often limited to introspective methods (Marks, 1973) and cleverly designed behavioural experiments (Kosslyn, Ball, & Reiser, 1978). Conducting research on mental imagery in animal models is even less controllable than in the human. Furthermore, neuroimaging methods hitherto did not provide a sufficient level of resolution to reconstruct mental images. This thesis, however, made use of state-of-the-art advancements in human neuroimaging methods to answer the posed research question. These first steps in reconstructing rather than only classifying a mental image are seminal to the advancement of cognitive neuroscience.

The successful decoding of complex mental images (like letter shapes) enables a new generation of BCIs. The practical implementations are (still) problematic as ultra-high field imaging is expensive and not wide-spread. By improving decoding accuracies further, however, there are potentially huge benefits to BCI users in that the communication with the computer and, thus, the interaction with the environment could become much more natural and broad. Feature-level fMRI based BCIs could accomplish something that no BCI is capable of at the moment: if the mapping of visual features is understood well enough and the top-down control of these features during imagery is of a sufficient level – both presented as a proof-of-concept in [Chapter 4](#) – the BCI user could freely generate decodable mental images without any restriction

*Building
natural brain-
computer
interfaces*

to a pre-made alphabet or other arbitrary communication protocol.

The studies presented in this thesis are part of a pioneering field in cognitive neuroscience, both on content-level as well as with respect to the methods involved. In the following paragraphs, the difficulties and considerations on the track towards high-field neuroimaging of mental imagery are described.

*The ultra-high
resolution
paradox*

When faced with functional fMRI images at submillimetre spatial resolution, the first reaction of cognitive neuroscientists often is amazement. Rightly so, as this unprecedented level of detail allows for a qualitatively new insight into the functional architecture of the human brain. However, this spatial resolution has a price besides the actual costs of state-of-the-art fMRI machines. Procedural as well as computational pipelines need to be adapted, if not redesigned, to fit the vast amounts of data that are produced by ultra-high field scanners. When fMRI was first used in cognitive neuroscience, typical voxel sizes were around 3 mm isotropic (Engel, Glover, & Wandell, 1997) which already posed computational problems for computer hardware and software at that time. Running experiments with submillimetre resolution fMRI data can produce datasets that are about one order of magnitude larger than conventional fMRI datasets (e. g., the volume of one 3 mm isotropic voxel is 27 mm^3 while the volume of one 0.8 mm isotropic voxel [as used in Chapter 3 and Chapter 4] is ca. 0.5 mm^3 which approximately corresponds to an increase of the number of voxels by a factor 53, given the coverage would be the same). Although computing power increased alongside this development (Schaller, 1997), vast amounts of data constitute a methodological challenge beyond compute time. It takes considerable efforts to run, for instance, searchlight analyses on high-field data (see Chapter 2 and

Chapter 3). Multicore parallelisation and computations on graphical processors (GPUs) have to be used in order to cope with the computational demands. These difficulties make us face a trade-off between a ‘zoom’ into a brain area of interest and the ability to investigate cognitive phenomena on a more global, whole-brain scale. Overall, this calls for well-planned methodological decisions to choose the neuroimaging method and parameters that fit the research questions best.

State-of-the-art ultra-high field fMRI is on the cusp of reaching functional resolutions that allow one to map columnar organisations in the human brain (De Martino et al., 2015; Yacoub, Shmuel, Logothetis, & Uğurbil, 2007; Zimmermann et al., 2011). In all studies presented in this thesis, especially in Chapter 2, we leveraged this new possibility to map feature representations in cortical areas that were not observable before. While sequences were optimised for acquisition speed, spatial resolution, and the covered field-of-view, they begin to reach a level of detail where the physiological underpinnings of the blood-oxygen level dependent (BOLD) signal start to matter. In this thesis, we employed gradient echo (GRE) echo-planar imaging (EPI) sequences to acquire data at submillimetre resolution. There is an ongoing debate about the origin of BOLD signals during GRE sequences and their spatio-temporal point spread function (Shmuel, Yacoub, Chaimow, Logothetis, & Uğurbil, 2007). Several studies (De Martino, Zimmermann, et al., 2013; Uğurbil et al., 2003; Uludağ, Müller-Bierl, & Uğurbil, 2009; Yacoub et al., 2007) suggest the use of spin-echo based sequences like GRASE (gradient- and spin-echo; Feinberg, Harel, Ramanna, Uğurbil, and Yacoub, 2008) to diminish the influence of large draining vessels in the hemodynamic signal and boost the sensitivity for blood oxygenation changes in the microvasculature. While an increase of actual functional resolution (rather than a mere tech-

*The
neurovascular
frontier*

nical reduction of voxel size), by means of adjusted physiological sensitivities, is required for mapping the functional architecture of cortical areas, the losses in signal-to-noise ratio (SNR) and field-of-view coverage often make the usage of these sequences difficult. Indeed, great efforts have to be made in order to reach high-resolution whole-brain coverage – even with gradient echo sequences – that run with acceptable volume repetition time (TR) and within specific absorption rate (SAR) limits. In a pilot experiment for the study in [Chapter 2](#), we used a 3D-GRASE sequence but ran into problems positioning the rather small slab onto bilateral hMT in every subject. Further development of fMRI sequences is ongoing (Kemper, De Martino, Vu, et al., 2015; Kemper, De Martino, Yacoub, & Goebel, 2015) and, for now, we did not yet reach an absolute limit of actual functional resolution of fMRI data. Despite contrast-to-noise ratio (CNR) and SAR considerations, the neurovascular nature of fMRI data will, however, result in an absolute functional resolution maximum (compared to electrophysiological neuroimaging methods) that is determined by the BOLD point spread function (Polimeni, Fischl, Greve, & Wald, 2010; Shmuel et al., 2007) and cannot be overcome by increasing the (technical) spatial resolution.

*The private
nature of
imagery*

Mental imagery is an inherently private phenomenon. It constitutes a cognitive process that is particularly difficult to control in an experimental setting. While this holds true for other complex cognitive domains like social interaction or planning, mental imagery still constitutes a special category: Humans perform mental imagery often without any behavioural outcome or proxy consequence, but mostly as an epiphenomenon or means throughout a reasoning or creative process. Some professions – visual artists, engineers, scientists – require a more streamlined form of mental imagery which made imagery research take interest in these

groups (Blazhenkova & Kozhevnikov, 2010; Ferguson, 1977). Nevertheless, even in these populations mental imagery eludes many common methods of experimental operationalisation and control. If the general principles behind visual mental imagery are to be found, it is all the more important to keep experimental control at a possible maximum by using training and precise instructions. In our imagery studies, it turned out to be crucial to extensively train our subjects in the mental imagery tasks before the actual experiment. Pilot studies resulted in useless data, when conducted with subjects that only relied on their spontaneous mental imagery capabilities. Instead, precise instructions about the mental imagery were sought after. Exploring and training the particular imagery tasks before the experiment resulted in more subjective confidence about the imagery performance and in significant outcomes.

This thesis, in many respects, contributes to cognitive neuroscience only the first steps towards a better understanding of the fine-grained organisation of human (visual) cortex, the nature of visual mental imagery, and towards enhanced natural BCIs. In the next paragraphs, possible future directions will be highlighted.

Homologies between visual neuroscience results in animal models and visual neuroscience in the human have been hypothesised and later verified time and time again. The in-vivo mapping of columnar functional organisations in the human became feasible with submillimetre resolution fMRI and led to the results presented in this thesis. There are still many exciting feature representations found in animal models, for instance facial feature mappings (Freiwald, Tsao, & Livingstone, 2009) or object feature mappings (Tanaka, 2003), that still need to be demonstrated in the human visual system. Furthermore, human representations could differ from the homologue animal representations. For instance, face-

*Towards
complex
feature-level
mapping*

selective areas in humans might be specific to the morphological properties of our own species. First steps that will benefit from increased functional resolution are being taken in this direction (Henriksson, Mur, & Kriegeskorte, 2015). More than that, however, it will become interesting to investigate feature-level organisations in the human visual system that cannot be observed in animal models. One such candidate could be the feature-level organisation of visual word form area (McCandliss, Cohen, & Dehaene, 2003), a region that is probably most pronounced in the human. Controversial findings with respect to the computational properties of this area (Rauschecker, Bowen, Parvizi, & Wandell, 2012) call for future experiments (Goebel, 2012) that could help to unravel the underlying functional representations. Finally, it will be interesting to explore feature-level representations beyond the human visual system. As mentioned in Chapter 1, the visual system lends itself well to neuroimaging research, however, studies in other cortical areas, such as the auditory system (De Martino et al., 2015; De Martino, Moerel, Van de Moortele, et al., 2013), have equally taken advantage of increased functional resolution.

The individual in imagery research

Generalisation is a main principle and driving force in neuroscience. Researchers try to find underlying principles that explain the bulk share of observed behaviour and label the remaining variance as error. While this is a serviceable approach in most of the cases and is reflected in the ‘work-horse’ methods of neuroscience – for instance the general linear model – fallacies occur when a possible effect of interest is labelled as unexplained (and therefore erroneous) variance. In the case of mental imagery, aspirations towards generalisation are at least partially moribund. The inter-individual differences that can be found during mental imagery tasks are huge and hitherto largely unexplained. Questionnaire data were collected to explain inter-individual differences

(Blazhenkova & Kozhevnikov, 2009), but surprisingly little neuroscientific evidence on this matter has been reported. Our own findings in [Chapter 3](#) encouraged us to publish a data paper with the individual searchlight accuracy maps, to allow other researchers to compare their findings during mental imagery tasks to ours. Although we acquired questionnaire data and the reported imagery strategies in our mental imagery studies of [Chapter 3](#) and [Chapter 4](#), our approach was not sufficient to holistically account for inter-individual differences. There is a need for more systematic investigations of the inter-individual differences during mental imagery. More knowledge about these factors could be of great benefit, both for the goal of unravelling the general functional principles of mental imagery as well as for the goal of closing the gap between the private nature of a mental image and the external observation thereof.

This thesis cannot be complete without mentioning the controversial topic of brain-reading. In [Chapter 3](#) and [Chapter 4](#), we demonstrated the possibility of decoding and even reconstructing mental imagery with compelling accuracy and clarity. When results like these are communicated to the broader public, oftentimes discussions about utopian or dystopian scenarios arise. In these discussions, the possibility of decoding what a person imagines goes hand in hand with the concept of brain-reading, that is, being able to read a person's thoughts. It may take a five-year PhD trajectory to realise that we could not be farther from this scenario. The main component of successful mental imagery experiments is compliance of subjects. If the subjects in an experiment are not trained well (as described above), meticulously follow task instructions, and completely comply with exhaustive experimental line-ups, there is absolutely no possibility to decode the mental images or even general thoughts they might have. The spatial and, more

*The
non-existing
brain-reading
dystopia*

importantly, accompanying temporal resolution and CNR of currently available neuroimaging methods (including ultra-high field fMRI) would need to be much higher to enable 'real' brain-reading. Even then, an incoherent subject would arguably still destroy any chances of recovering mental images against their will.

There is, however, a special group of subjects with maximal compliance: BCI users. If other means of communication are rendered impossible for an individual, the utility of a brain-reading device increases dramatically. BCI users are willing to train their mental imagery capabilities and could yield sufficient control over a new generation of BCIs. By leveraging feature-level representations that enable the decoding and reconstruction of freely chosen mental images, we can create a utopian version of brain-reading. In this scenario, BCI users – and possibly more people – can use their mental images and thoughts to communicate with their environment in novel ways, however, exclusively when they want to do so.

REFERENCES

- Blazhenkova, O. & Kozhevnikov, M. (2009, July). The new object-spatial-verbal cognitive style model: Theory and measurement. *Applied Cognitive Psychology*, 23(5), 638–663. doi:[10.1002/acp.1473](https://doi.org/10.1002/acp.1473)
- Blazhenkova, O. & Kozhevnikov, M. (2010, December). Visual-object ability: a new dimension of non-verbal intelligence. *Cognition*, 117(3), 276–301. doi:[10.1016/j.cognition.2010.08.021](https://doi.org/10.1016/j.cognition.2010.08.021)
- De Martino, F., Moerel, M., Uğurbil, K., Goebel, R., Yacoub, E., & Formisano, E. (2015). Frequency preference and attention effects across cortical depths in the human primary auditory cortex. *Proceedings of the National Academy of Sciences*, 112(52), 16036–16041. doi:[10.1073/pnas.1507552112](https://doi.org/10.1073/pnas.1507552112)
- De Martino, F., Moerel, M., Van de Moortele, P.-F., Uğurbil, K., Goebel, R., Yacoub, E., & Formisano, E. (2013). Spatial organization of frequency preference and selectivity in the human inferior colliculus. *Nat Commun*, 4, 1386. doi:[10.1038/ncomms2379](https://doi.org/10.1038/ncomms2379). arXiv: [NIHMS150003](https://arxiv.org/abs/NIHMS150003)
- De Martino, F., Zimmermann, J., Muckli, L., Uğurbil, K., Yacoub, E., & Goebel, R. (2013, January). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. *PloS one*, 8(3), e60514. doi:[10.1371/journal.pone.0060514](https://doi.org/10.1371/journal.pone.0060514)
- DeAngelis, G. C. & Newsome, W. T. (1999, February). Organization of disparity-selective neurons in macaque area MT. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 19(4), 1398–415. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9952417>
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997, March). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral cortex (New York, N.Y. : 1991)*, 7(2), 181–92. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9087826>
- Feinberg, D. A., Harel, N., Ramanna, S., Uğurbil, K., & Yacoub, E. (2008). Submillimeter Single-shot 3D GRASE with Inner Volume Selection for T2 weighted fMRI applications at 7 Tesla 1. *Proc. Intl. Soc. Mag. Reson. Med.* 16, 37(4), 2373.
- Ferguson, E. S. (1977, August). The Mind's Eye: Nonverbal Thought in Technology. *Science (New York, N.Y.)* 197(4306), 827–36. doi:[10.1126/science.197.4306.827](https://doi.org/10.1126/science.197.4306.827)

- Freiwald, W. A., Tsao, D. Y., & Livingstone, M. S. (2009, September). A face feature space in the macaque temporal lobe. *Nature neuroscience*, 12(9), 1187–96. doi:[10.1038/nn.2363](https://doi.org/10.1038/nn.2363)
- Gilbert, C. D. & Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. *Neuron*, 54(5), 677–696. doi:[10.1016/j.neuron.2007.05.019](https://doi.org/10.1016/j.neuron.2007.05.019)
- Goebel, R. (2012, June). Position coding in the visual word form area. *Proceedings of the National Academy of Sciences of the United States of America*, 109(24), 9226–7. doi:[10.1073/pnas.1206648109](https://doi.org/10.1073/pnas.1206648109)
- Goncalves, N. R., Ban, H., Sanchez-Panchuelo, R. M., Francis, S. T., Schluppeck, D., & Welchman, A. E. (2015). 7 Tesla fMRI Reveals Systematic Functional Organization for Binocular Disparity in Dorsal Visual Cortex. *Journal of Neuroscience*, 35(7), 3056–3072. doi:[10.1523/JNEUROSCI.3047-14.2015](https://doi.org/10.1523/JNEUROSCI.3047-14.2015)
- Henriksson, L., Mur, M., & Kriegeskorte, N. (2015). Faciotopy-A face-feature map with face-like topology in the human occipital face area. *Cortex*, 72, 156–167. doi:[10.1016/j.cortex.2015.06.030](https://doi.org/10.1016/j.cortex.2015.06.030)
- Kemper, V. G., De Martino, F., Vu, A. T., Poser, B. A., Feinberg, D. A., Goebel, R., & Yacoub, E. (2015, May). Sub-millimeter T₂-weighted fMRI at 7 T: Comparison of 3D-GRASE and 2D SE-EPI. *Frontiers in Neuroscience*, 9(APR), 1–14. doi:[10.3389/fnins.2015.00163](https://doi.org/10.3389/fnins.2015.00163)
- Kemper, V. G., De Martino, F., Yacoub, E., & Goebel, R. (2015, September). Variable flip angle 3D-GRASE for high resolution fMRI at 7 tesla. doi:[10.1002/mrm.25979](https://doi.org/10.1002/mrm.25979)
- Kok, P., Bains, L. J., van Mourik, T., Norris, D. G., & de Lange, F. P. (2016). Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. *Current Biology*, 1–6. doi:[10.1016/j.cub.2015.12.038](https://doi.org/10.1016/j.cub.2015.12.038)
- Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: evidence from studies of image scanning. *Journal of experimental psychology. Human perception and performance*, 4(1), 47–60. doi:[10.1037/0096-1523.4.1.47](https://doi.org/10.1037/0096-1523.4.1.47)
- Marks, D. F. (1973, February). Visual imagery differences in the recall of pictures. *British journal of psychology (London, England : 1953)*, 64(1), 17–24. doi:[10.1111/j.2044-8295.1973.tb01322.x](https://doi.org/10.1111/j.2044-8295.1973.tb01322.x)
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003, July). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, 7(7), 293–299. doi:[10.1016/S1364-6613\(03\)00134-7](https://doi.org/10.1016/S1364-6613(03)00134-7)

- Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Uğurbil, K., ... Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology*, 25(20), 2690–2695. doi:10.1016/j.cub.2015.08.057
- Nasr, S., Polimeni, J. R., & Tootell, R. B. H. (2016). Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. *Journal of Neuroscience*, 36(6), in press. doi:10.1523/JNEUROSCI.3518-15.2016
- Polimeni, J. R., Fischl, B., Greve, D. N., & Wald, L. L. (2010, October). Laminal analysis of 7T BOLD using an imposed spatial activation pattern in human V1. *NeuroImage*, 52(4), 1334–1346. doi:10.1016/j.neuroimage.2010.05.005
- Rauschecker, A. M., Bowen, R. F., Parvizi, J., & Wandell, B. A. (2012, June). Position sensitivity in the visual word form area. *Proceedings of the National Academy of Sciences of the United States of America*, 109(24), E1568–77. doi:10.1073/pnas.1121304109
- Schaller, R. R. (1997). Moore's law: past, present and future. *Spectrum, IEEE*, 34(6), 52–59. doi:10.1109/6.591665
- Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N. K., & Uğurbil, K. (2007, April). Spatio-temporal point-spread function of fMRI signal in human gray matter at 7 Tesla. *NeuroImage*, 35(2), 539–52. doi:10.1016/j.neuroimage.2006.12.030
- Tanaka, K. (2003, January). Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cerebral cortex (New York, N.Y. : 1991)*, 13(1), 90–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12466220>
- Uğurbil, K., Adriany, G., Andersen, P., Chen, W., Garwood, M., Gruetter, R., ... Zhu, X.-H. (2003, December). Ultrahigh field magnetic resonance imaging and spectroscopy. *Magnetic Resonance Imaging*, 21(10), 1263–1281. doi:10.1016/j.mri.2003.08.027
- Uludağ, K., Müller-Bierl, B., & Uğurbil, K. (2009, October). An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. *NeuroImage*, 48(1), 150–65. doi:10.1016/j.neuroimage.2009.05.051
- Yacoub, E., Harel, N., & Uğurbil, K. (2008, July). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 105(30), 10607–12. doi:10.1073/pnas.0804110105

- Yacoub, E., Shmuel, A., Logothetis, N. K., & Ugurbil, K. (2007, October). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *NeuroImage*, 37(4), 1161–77. doi:[10.1016/j.neuroimage.2007.05.020](https://doi.org/10.1016/j.neuroimage.2007.05.020)
- Zimmermann, J., Goebel, R., De Martino, F., Van de Moortele, P.-F., Feinberg, D. A., Adriany, G., . . . Yacoub, E. (2011, December). Mapping the Organization of Axis of Motion Selective Features in Human Area MT Using High-Field fMRI. *PLoS ONE*, 6(12), e28716. doi:[10.1371/journal.pone.0028716](https://doi.org/10.1371/journal.pone.0028716)

SUMMARY

This thesis sheds light on the functional architecture of several cortical areas in the human visual system during visual perception and imagery. It employed high-resolution 7 Tesla functional magnetic resonance imaging (fMRI) for both mapping and decoding of brain activation patterns within human visual cortex. Depth perception evoked by binocular disparity including multiple visual feature encodings, visual imagery of motion as well as letters, and top-down influences on the visual system exploited by mental imagery have been investigated. A general introduction to the subject matter is presented in chapter one.

In [Chapter 2](#) (“Overlapping feature representations in human visual cortex: Binocular disparity and axis-of-motion tuning mapped in human area MT using ultra-high field fMRI”), depth perception induced by binocular disparity was investigated. This led to a deeper understanding of the functional architecture within the human middle temporal area (hMT) with respect to its computational properties and encodings. Exploiting submillimetre functional resolution ultra-high field fMRI, binocular disparity and direction-of-motion tunings in hMT were examined. Systematic fine-grained disparity tunings, alongside axis-of-motion tunings were found. This study was the first to demonstrate overlapping tuning maps for two distinct visual features, i.e. binocular disparity and direction-of-motion, in an extrastriate area of human visual cortex using a fully crossed and controlled experimental design. This shed light on how the brain maps multiple basic visual features within one cortical area.

Chapter 3 (“Decoding the direction of imagined visual motion using 7T ultra-high field fMRI”) continued this line of thought by examining not only perceived, but also merely imagined visual content within the same hierarchical structures of the visual system. The direction of imagined visual motion was successfully decoded from primary visual areas on a single-subject level. Top-down influences during imagery led to remarkably similar activation in early visual cortex as perception. The application of multi-voxel pattern analysis to high-resolution functional data of twelve subjects demonstrated that imagery of visual motion activated the earliest levels of the visual hierarchy and that the extent and location of the activation varied between subjects. The approach enabled classification not only of complex imagery, but also of actual content, in that the direction of imagined motion, out of four options, was successfully identified in most subjects and with high accuracies. These high-accuracy findings not only shed light on the constituents of mental imagery, but also on the feasibility of decoding specific sub-categorical imagery content from brain imaging data.

Finally, in **Chapter 4** (“Reconstructing and decoding imagined letters from the visual system using ultra-high field fMRI”), the previous results were taken one step further and applied to enable visual reconstruction of concrete specific imaging content. For the first time, the successful reconstructions of visual field images (visualising recognisable content of a complex mental image) was achieved on a single trial level. Imagined letter shapes were reconstructed and decoded from submillimetre resolution fMRI data of two subjects baring striking resemblance to the presented letter shapes, especially when focusing on the earliest visual areas. This demonstrated significant overlap between reconstructed visual field images of perception data, imagery data, and the ac-

tually presented stimuli. Furthermore, successful classification of different letter features (edgy versus curvy) based on data from visual word-form area using multi-voxel pattern analysis was accomplished. This proof of concept could be the first step towards more elaborate studies on visual field image reconstruction of imagined everyday stimuli and towards the development of more natural and directly content-based brain-computer-interfaces.

A discussion of all findings is presented in chapter five. In summary, this thesis provided new evidence on the fine-grained organisation of visual areas in the human brain during visual perception and during mental imagery, employing state-of-the-art ultra-high resolution functional imaging.

VALORISATION

The work presented in this thesis describes neuroscientific results on the fine-grained functional organisation of brain areas in the human visual system during perception and imagery. These results are quite hard to “valorise” but rather fundamental research findings with an inherent value.

In the discussions of [Chapter 3](#) and [Chapter 4](#), as well as in the General Discussion in [Chapter 5](#) the possibilities of building advanced brain-computer interfaces (BCIs) that would make use of the results at hand are mentioned. If one is very eager to valorise these studies one could, indeed, try to put a 7 Tesla functional magnetic resonance imaging (fMRI) scanner into the living room of a potential BCI user and enable her to steer a computer cursor using motion imagery (e. g., a centimetre in one of four directions every 20 seconds) or imagine a letter to spell (one letter out of ‘H’, ‘T’, ‘S’, or ‘C’ every 15 seconds). This would be an extremely costly way of steering a computer cursor or typing letters, considering the costs of such an fMRI scanner. These costs and the required environment for an ultra-high field fMRI scanner make a normal usage in one’s the living room frankly impossible. Furthermore, the rather low decoding accuracies for some subjects described in [Chapter 3](#) and [Chapter 4](#) would require many trial repetitions increasing the confidence of the BCI system to acceptable levels. The rather uncomfortable BCI user experience and the need for highly qualified personnel to operate the fMRI scanner make this BCI even more impractical. Taking into account the costs (and practical is-

*Valorisation
as BCIs*

sues) of an ultra-high field fMRI scanner, the bitrates (with which the BCI can translate intentions into actions), and decoding accuracies, this does not offer a competitive solution when compared to readily available BCI system like an electroencephalography (EEG) P300 speller. The direct translation of the results presented in this thesis into an fMRI-based BCI is, thus, not a very promising valorisation.

*Actual
valorisation
as BCIs*

The results presented in this thesis are, however, potentially useful in a BCI context. Neuroimaging research on mental imagery can potentially enable a future BCI system, based on a new neuroimaging technique to locate and decode the neural correlates of mental imagery in a more guided and straightforward way. It is conceivable that, e. g., more invasive neuroimaging techniques like electrocorticography (ECoG) or multielectrode arrays (MEAs), or even not yet invented (potentially) non-invasive high-resolution neuroimaging techniques would make use of the feature representations described in the studies of this thesis. The information about the fine-grained cortical organisations of feature maps during imagery could guide the development of such BCIs. Moreover, insights into inter-individual differences in imagery strategies as well as brain activations during mental imagery – as discussed throughout this thesis – are important to successfully implement any BCI system and could, therefore, also improve the performance of other non-fMRI-based BCIs.

Finally, there are medical conditions in which high-field fMRI-based BCIs can become important. Some locked-in syndrome patients (e. g., some patients with amyotrophic lateral sclerosis (ALS)) cannot successfully use BCIs based on other neuroimaging methods (e. g., EEG or functional near-infrared spectroscopy (fNIRS)). In such extreme situations when all other means of communication are lost and the quality of life is severely affected, the above-

mentioned impracticalities lose their importance. The ability to communicate with one's environment cannot be matched with costs for technical equipment. The results presented in this thesis could be used in such rare cases not only to allow the patient to communicate again but could also outperform BCI results based on conventional fMRI data (e. g., 3 Tesla).

Besides the inherent value of basic research results presented in this thesis I see the main value of the work I did over the last few years not in potential (BCI) products. I see value in the methodological and practical scientific knowledge that I gained during my work and tried to share with fellow researchers. The photo on the cover of this thesis shows a lonely sculler on a (seemingly wide and) empty ocean. It depicts something that immediately evokes some form of imagery in the viewer ('Who is this person and where does she come from?' or 'How long is she already sculling?') and, therefore, fits the title of my thesis. However, the scene also resembled a feeling one can easily get when confronted with neuroscientific research projects. Especially, when new methods are involved (like ultra-high field fMRI) or largely uncharted scientific territory is entered (like neuroimaging of mental imagery) the problems that arise during research do not seem to end. From the beginning of my PhD studies I could rely on the knowledge and help of others and I continued to help others in turn as soon as I started solving new problems arising from my own work. While there are many talented bright scientists working in cognitive neuroscience the task at hand – that is, to unravel the functional working principles of the human brain – demands the sharing of knowledge, not only in the form of scientific publications, but also in very practical terms. Wherever I could, I tried to spare time for the dissemination of, e. g., source code (contributions to the experiment software PsychoPy, Import/Export rou-

*Actual
valorisation*

tines for BrainVoyager file formats into Python), manuals (Eye-tracking on the 7T system, 7T segmentation guides, MNI space transformation guides to convert fMRI maps for data sharing platforms), or hands-on help. I am very happy to have worked in an environment where more people acted like I did, as the sharing of practical knowledge is crucial to the success of scientific projects. By making our work available in an accessible way (through open-access publications or by releasing source code and manuals under open source/libre licenses) and through collaborative efforts (like Wiki documentations or contributing to existing open source projects), we create immense value for the scientific process and, thereby, society in general.

ACKNOWLEDGEMENTS

It makes me very happy that I was able to work on this thesis not only on my own, but with the great help and good company of many people around me. Therefore, I would like to take the opportunity to thank some of these people.

I would probably not be in Maastricht without *Rainer*, who invited me first for an internship and later for my Diploma thesis and the PhD position. It has always been a great inspiration to work with you and experience your enthusiasm about technology. *Bettina*, you welcomed me in Maastricht and supported me throughout my Diploma thesis and the beginning of my PhD project. It was a lot of fun to explore the fNIRS world together with you and tap my nose into consortium meetings before I really knew what they were. Thank you! *Martin*, it was always great fun and help to meet with you (now my eyes are literally on the prize!) and your supportive and encouraging attitude got me through the tougher times of my PhD. I will actually really miss our Tuesday meetings and probably my English will get worse again.

I guess that without *Valentin* ('Isch hab Valentin!'), *Giancarlo*, and *Jan*, none of my projects would have been possible. It is not only scanning or some source code snippet – you guys were a great help over the last years, and – more than that – made work so much more fun. Thank you!

Helen, thanks for standing behind Franzi and me, metaphorically and literally! It was great fun to be paranymphs with you (and to make the best PhD movie ever; there should be Oscars after all) and it is wonderful to have Valentin and you as my support.

Mario, Arne, Valentin, Martin, Joel, Judith, and Sven, I enjoyed working and, moreover, having nice conversations with every single one of you!

Lars H., Dimo, Roy, and Bert, thanks for helping out so often with things that (in some cases) I will never fully understand. . .

I was also lucky to have had great office mates over the years: *Inge, Geraldine, Anke, and Mario*, I would not want to miss all the small chats with you that made for a really relaxed office atmosphere.

Christl, Riny, Eva, what should I do without you? What should anybody in CN do without you? You did not only help out with so many small and bigger issues, you did it in a way which made me walk out of your offices with a big smile.

Faruk, Carmine, Mario M., I loved the coffee corner talks about some new programming project, the next internet meme, or other geek stuff. You are great!

Arko, please consider a second career as a TV show host! I would watch every episode.

Alex and Teresa, thank you for welcoming us to Maastricht and advising us throughout the years. It was a pleasure to play the (good) Sinterklaas for Mathilda, Rosalie, and Oskar (who will hopefully not be flabbergasted if they read this one day).

Federico and Kâmil, thanks for the fun chats and the occasional opinionated argument. I actually really like that.

All the participants in my experiments, thanks for enduring long scans and (possibly) exhausting training sessions!

The whole CN department, you guys make for such a great atmosphere at work! I enjoy the conversations with every single one of you!

All the BrainInnovation folks, it is great to work with you! And I love it that all you guys know what a Raspberry Pi is.

Chris, Lotty, Esther, and Verena, thanks for all your support and chats over the years.

The Banditos team, you are probably the most charming Baristas I will ever meet.

Outside university, there are more people who supported me over the last years:

Rik, thanks for showing me that there is more to dare in life than I often thought. It has always been a great (and often crazy) time with you. The last years would have been definitely more boring without you! And never forget: Things are often more complicated than you thought ;)!!!

I will really miss the Looiersgracht Kommune in the next year: *Rik, Catharina, Christoph, and Nora*, you are the most wonderful and cheerful neighbors one could wish for!

Ulf, we know each other now for such a long time. I am so happy, that we manage to stay in close contact and I cherish every single phone call and meeting with you. You are a truly good friend!

Steene and *Lars*, thanks for being and staying good friends!

Patrick, thanks for the great conversations in Trier and later during our world-wide encounters. It is always an inspiration to talk to you and work with you.

All people of the TZ in Hamburg and the Lüneburger Heide that *Franzi* and me met over the years, you were a great support and gave me so many different perspectives on life.

And now on a more personal note:

Conny, Hannes, Resi, Johanna und *Marie*, ich freue mich wahnsinnig, dass Ihr in den letzten Monaten Daumen gedrückt habt und mit mir jetzt feiert! Ihr seid wirklich spitze und es ist toll sich in Nördlingen zuhause fühlen zu können.

Echen, Oma, Susanne und *Marianne*, es ist wunderbar Euch alle schon bei so vielen wichtigen Schritten in meinem Leben dabei gehabt zu haben und Eure Unterstützung zu spüren.

Martin, Claudia, Tobias und *Sebastian*, jeder Besuch von und bei Euch macht so viel Spaß und mischt den Alltag mal ein bisschen durch. Ich danke Euch für Eure offenen Ohren und Eure Herzlichkeit und Fröhlichkeit. *Martin*, ich danke Dir für all die langen abendlichen Gespräche! Durch Dich habe ich mein Interesse an Technologie entdeckt – Du bist ein wirklich wichtiger Mensch für mich und ein toller Patenonkel!

Liebe *Mama*, lieber *Papa*, alles, was ich bisher geschafft habe, ist nur möglich gewesen, weil Ihr mir einen so unglaublich guten Start in mein Leben gegeben habt! Ihr wart und seid wundervolle Eltern und ich bin Euch sehr dankbar für all die Unterstützung und Liebe, die ich durch Euch erfahre. Danke!

Franziska, das Leben mit Dir ist wunderschön. Ich kann mein Glück, Dich gefunden zu haben, kaum fassen und bin zutiefst dankbar für jeden Tag, den ich mit Dir verbringen darf. Die letzten Jahre haben sowohl mich als auch Dich manchmal auf die Probe gestellt, aber an uns zusammen konnte nie etwas rütteln. Deine Unterstützung in den letzten Jahren ist mit Worten nicht zu beschreiben. Ich danke Dir aus tiefstem Herzen und freue mich auf alle Schritte, die wir zusammen gehen werden! Und es wird immer noch schöner. . .

CURRICULUM VITÆ

Thomas Christoph Emmerling was born on September 23rd 1986 in Haan (Rheinland), Germany. His first computer was a Commodore 64. He obtained his High School Diploma in 2005 at the Ludwig-Meyn-Schule in Uetersen, Germany, and worked as an audio engineer at the Thalia Theater in Hamburg, Germany, during a voluntary social year in 2006. From September 2006, Thomas studied Psychology and Computer Science (minor) at the University of Trier, Germany, and graduated with a Diploma (B.Sc. and M.Sc.) in 2011. During his studies he worked as a student assistant in the EEG lab at the University of Trier under Dr. Ewald Naumann and became a fellow of the German National Merit Foundation in 2007. For his Diploma thesis he developed a hybrid EEG-fNIRS brain-computer interface under the supervision of Prof. Dr. Rainer Goebel and Dr. Bettina Sorger. In September 2011, he started his PhD project as part of the ERC grant “Columnar Code Cracking” under the supervision of Prof. Dr. Rainer Goebel and Dr. Martin Frost at the Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University. From March 2016, Thomas will be working as a software engineer at Brain Innovation B.V., Maastricht.

PUBLICATIONS

PEER REVIEWED JOURNAL ARTICLES

Emmerling, T. C., Senden, M., Frost, M. A., & Goebel, R. (*in preparation*). Reconstructing and decoding imagined letters from the visual system using ultra-high field fMRI.

Kemper, V. G., De Martino, F., **Emmerling, T. C.**, Yacoub, E., & Goebel, R. (*in preparation*). 9.4T anatomical and functional investigation of human primary and secondary visual cortex.

Emmerling, T. C., Valente, G., Zimmermann, J., Kemper, V. G., Frost, M., & Goebel, R. (*submitted*). Overlapping feature representations in human visual cortex: Binocular disparity and axis-of-motion tuning mapped in human area MT using ultra-high field fMRI.

Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A., & Goebel, R. (2016). Time-resolved searchlight analysis of imagined visual motion using 7T ultra-high field fMRI: Data on interindividual differences. *Data in Brief*, 7, 468–471.

<http://doi.org/10.1016/j.dib.2016.02.071>

Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A., & Goebel, R. (2016). Decoding the direction of imagined visual motion using 7T ultra-high field fMRI. *NeuroImage*, 125, 61–73.

<http://doi.org/10.1016/j.neuroimage.2015.10.022>

BOOK CHAPTERS

Guger, C., Sorger, B., Noirhomme, Q., Naci, L., [...] **Emmerling, T. C.**, [...] & Kübler, A. (2014). *Brain-computer interfaces for assessment and communication in disorders of consciousness*. In G. Naik & Y. Guo (Eds.) *Emerging theory and practice in neuroprosthetics* (pp. 181–214). Hershey, PA: IGI Global.

INVITED TALKS

Emmerling, T. C., Sorger, B., Zimmermann, J., & Goebel, R. (2013). Decoding imagined visual motion in different directions using conventional 3 Tesla and high spatial resolution 7 Tesla fMRI. Talk presented at *Neuroscience 2013* in San Diego (CA).

Emmerling, T. C. (2013). Combining fNIRS & EEG: Principles and applications. Talk presented in the *ABC brain imaging workshop*, Maastricht, April 2013.

Emmerling, T. C. (2012). Combined EEG and fNIRS BCI in Humans. Talk presented in the educational course “Brain-Computer Interfaces and Real-Time fMRI” at the 18th *Annual Meeting of the Organization for Human Brain Mapping* in Beijing, China, June 2012.

CONFERENCE PRESENTATIONS

Emmerling, T. C., Zimmermann, J., Frost, M. A., & Goebel, R. (2014). Disparity tuning in human area MT measured using

sub-millimeter resolution 7 Tesla fMRI. Poster presented at *Neuroscience 2014* in Washington, D.C.

Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A., & Goebel, R. (2014). Decoding imagined visual motion in different directions using high spatial resolution 7 Tesla fMRI. Poster presented at the 20th *Annual Meeting of the Organization for Human Brain Mapping* in Hamburg, Germany.

Emmerling, T. C., Sorger, B., Goebel, R. (2012). Decoding the direction of visual motion in perception and imagery using multi-voxel pattern analysis. Poster presented at the 18th *Annual Meeting of the Organization for Human Brain Mapping* in Beijing, China.

Müller, F.-N., **Emmerling, T. C.**, Naumann, E., & Britz, P. (2011). Öhman paradigm: An approach to explaining inconsistent results in replication studies by integrating eyetracking into EEG experiments. Poster presented at the 53rd *Tagung experimentell arbeitender Psychologen*, Halle (Saale), Germany.

Strelzyk, F., Hermes, M., **Emmerling, T. C.**, Bergmann, R., Yakovlev-Leyendecker, O., Naumann, E., & Schächinger, H. (2010). Schnelle, nicht-genomische Effekte von Cortisol auf die Hirnperfusion von 28 gesunden Männern - eine Continuous Arterial Spin Labeling (CASL) Studie. Poster presented at the 36th "*Psychologie und Gehirn*" in Greifswald, Germany.

Strelzyk, F., Bertsch, K., **Emmerling, T. C.**, Siebenbrock, A., & Naumann, E. (2009). Experimental ERP effects do not coincide with the largest amplitudes of components in a visual search task. Poster presented at the 49th *Annual Meeting of the Society for Psychophysiological Research* in Berlin, Germany.

Siebenbrock, A., **Emmerling, T. C.**, Fechtner, J., Naumann, E., & Sollers, J. (2009). Herzratenvariabilität, Aufmerksamkeits- und Emotionsregulation. Poster presented at the 35th "*Psychologie und Gehirn*" in Leipzig, Germany.

