Fachbereich Erziehungswissenschaft und Psychologie

der Freien Universität Berlin

AN INTERPLAY OF FEEDFORWARD AND FEEDBACK SIGNALS SUPPORTING VISUAL COGNITION

Dissertation

zur Erlangung des akademischen Grades

Doktor(in) der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von

M. Sc.

Iamshchinina, Polina

Berlin, 2021

Erstgutachter: Prof. Dr. Radoslaw Martin Cichy

Zweitgutachter: Prof. Dr. Felix Blankenburg

Tag der Disputation: 06.04.2022

CONTENTS

1 INTRODUCTION

It is not enough to react stereotypically to stimuli to enable adaptive and intelligent behaviour in a complex environment. Instead, from the large stream of information reaching the organism at any moment, the relevant information must be prioritised. A large body of research demonstrated that a wide range of complex cognitive functions such as goal-directed attention as well as perception and working memory rely on internally-generated signals modulating sensory input (Saenz, Buracas & Boynton, 2002; Self et al, 2013; Zhaoping, 2019). Despite the substantial role of these modulatory influences in cognitive functioning, to date little is known about how they are implemented in the brain and what their contents constitute (Bastos et al., 2012; Pozzi, Bothe & Roelfsema, 2018; Lillicrap et al., 2020).

The modulatory influences are supported by a complex net of recurrent connections in the brain: Abundant anatomical projections run not only from early sensory regions towards downstream brain areas (feedforward) but also in the opposite direction (feedback) (Felleman & van Essen. 1991), thereby enabling interareal crosstalk even in the absence of sensory input. The feedback projections tightly connect brain areas within the visual system (Michalareas et al., 2016). Studies of visual mental imagery and visual working memory confirm this tight connection functionally by showing a multi-level propagation of internally-generated representations from frontal eye fields and lateral occipital cortex all the way down to early visual areas (Ester et al., 2015; Lee, Kravitz & Baker, 2012). These studies provide compelling evidence that feedforward and feedback signals are concurrently represented in the same visual brain regions. However, the signal representations strongly overlap at the level at which they were measured. Therefore, it remains unclear how the visual brain separates feedforward and feedback signals, thus avoiding the mixture of the perceived and the imagined.

Feedback projections do not only connect brain areas within the visual system but are also known to closely connect brain regions in different sensory systems (Rockland & Ojima, 2003; Rockland & van Hoesen, 1994; Eckert et al., 2008; Beer et al., 2013). This brings up an interesting complementary question: how are visual representations activated by signals from other sensory systems? There is abundant evidence that representations in one sensory modality can be triggered by those from the other sensory modality: Natural sounds and spoken words elicit selective responses in early visual, fusiform and parahippocampal areas (Vetter, Smith & Muckli, 2014; O'Craven & Kanwisher, 2000), tactile mental imagery can be initiated by visually presented cue (Schmidt & Blankenburg, 2019), while visual instructions can initiate speech mental imagery (Lu et al., 2019). What mechanisms underlie this prompt transition of signals across sensory systems remains to be investigated.

Building on the previous studies, we pose **two critical questions** about feedback modulations coming from within the same or from a different sensory system, respectively: **(1) How are visual feedback representations not confused with visual feedforward input? (2) How are visual representations activated by feedback signals from another sensory system?** The next two chapters will describe state-of-the-art research addressing these two questions.

1.1 How are visual feedback representations not confused with visual feedforward input?

Research focusing on delineation of feedforward and feedback signals branched off in two directions, namely, the differentiation by: (1) frequency bands within which the signals are transmitted and (2) cortical laminae through which these signals pass. A series of studies led by Fries and colleagues (Fries at al., 2001; Fries, 2015) investigated how rhythmic synchronization subserves directional signal transmission. First, this research group demonstrated that perception of a visual stimulus enhances synchronization between V1 and V4 in the gamma-frequency band

in macaque brain (Bosman et al., 2012; Rohenkohl, Bosman & Fries, 2018). Next, they explored if there is a relation between frequency-band rhythms and directed (feedforward/feedback) anatomical projections in the primate visual system (Bastos et al., 2015; Mejias et al., 2016). They found that initially identified gamma-band rhythm carries signals in the feedforward direction, while alpha/beta band subserves signal transmission in the feedback direction. Finally, these correlations were also established between the synchronization bands estimated from MEG human data and directed anatomical projections based on macaque brain (Michalareas et al., 2016).

Building on the revealed specific role of frequency-band rhythms in transmission of feedforward and feedback signals, it is natural to assume that the experimental task manipulations of feedforward/feedback information streams in humans may reveal their differential representation in frequency space. However, the explicitly manipulated feedforward and feedback influences were found to be homogeneously distributed in frequency space when measured with EEG in humans (e.g., Xie, Kaiser & Cichy, 2020) or did not align with the frequency bands predicted from the animal data (Canales-Johnson et al., 2021). Overall, this series of studies revealed that specific frequency-band rhythms carry signals in feedforward/feedback directions in the visual system. However, an experimental manipulation of the feedforward/feedback signals has not yet revealed their separation by frequency-bands in humans, possibly due to insufficient signal sensitivity of the human neuroimaging methods compared to primate electrophysiology.

In contrast, explicit manipulation of the feedforward/feedback signals did reveal effective differentiation of these signals by cortical depth both in primates and humans (for a review see Self et al., 2019). In animal work, stimulus-induced activity in V1 was shown to first emerge in the middle laminae, whereas feedback activity induced by figure-ground segregation, maintenance in working memory and covert attention shifts emerged in the deep and superficial layers (van

Kerkoerle et al., 2017). The results of this functional study fully correspond to an early neuroanatomical finding (Rockland & Pandya, 1984): Abundant feedforward connections project predominantly to the middle layer, while sparse feedback connections terminate outside the middle layer (Felleman and Van Essen, 1991; Bastos et al., 2012; Barone et al., 2000). Furthermore, 7T fMRI in humans allowed for sufficient spatial resolution to test the differentiation of feedforward and feedback signals by cortical depth in humans. Consistent with the results of animal research (Self et al., 2013; Kerkoerle et al., 2017), feedback-induced activations in humans were identified in one of the outer grey matter depth bins (superficial or deep) using a diversity of paradigms and analysis pipelines (Muckli et al., 2015; Kok et al., 2016; Lawrence et al., 2018; Aitkens et al., 2021; Yu et al., 2019). Feedforward activations were spread over the grey matter with some predominance at the middle cortical depth (Muckli et al., 2015; Kok et al., 2016). Thus, the signal separation by cortical layer was established in animal research and the depth-dependent studies in humans generally correspond to the animal findings.

Despite the overall good alignment of the results across species, feedforward/feedback differentiation with 7T fMRI in humans faces several challenges. Firstly, previous studies estimated feedforward and feedback signals in isolation using separate tasks. Instead, feedforward and feedback signals should be estimated in close-to-equal conditions to show that the differences in cortical depth induced by these signals are not due to task-related effects. Additionally, feedforward and feedback streams naturally happen concurrently rather than sequentially and thus their spatial dynamics might simply differ when each signal is estimated in isolation (e.g., feedback signal estimated in the presence of sensory input vs. when eyes closed is likely to show different depth distribution). Secondly, the majority of the results in human studies are based on data acquired with conventional gradient-echo echo-planar imaging which despite having a high signal

sensitivity shows low spatial specificity and is susceptible to influences from large locally unspecific blood vessels distant of the activated cortical layer (Gagnon et al., 2016; Koopmans et al., 2010; Menon et al., 1995; Turner, 2002). The effect of these confounds can be diminished with careful experiment design, analysis and post-processing (Polimeni et al., 2010; Koopmans, Barth & Norris, 2010; Uludag & Blinder, 2018), but only to an extent.

Altogether, in order to explore how visual feedback representations are not confused with visual feedforward input, we conducted Study I (described below in more detail), where we aimed to overcome the limitation of isolated signal estimation in cortical depth. Furthermore, we conducted Study II to address the challenge of low spatial specificity when measuring cortical depth-dependent responses.

1.2 How are visual representations activated by feedback signals from another sensory system?

While the first critical question was focused on the separation of feedforward and feedback signals within the visual system, here we address the research investigating how signals from another sensory system initiate visual representations. Several neuroanatomical studies found a feedback-type connection from auditory to visual early cortices - either direct ones or via higher-order areas (Rockland & Ojima, 2003; Rockland & van Hoesen, 1994; Cappe & Barone, 2005). Next, a series of functional studies confirmed a direct link between the auditory signals and the content-specific responses propagating through the visual brain regions (Vetter et al., 2014; Vetter et al., 2020). It was found that natural sound categories (e.g., “human” and “inanimate”) can be decoded in sighted blindfolded participants from early visual areas, superior temporal gyrus and auditory cortices (Vetter et al., 2014). As of now, it is not entirely clear if this category information is relayed to early visual areas via multisensory regions such as superior temporal gyrus or via direct anatomical connections between the auditory and visual early areas. The presence of the

auditory-induced information in visual areas might be due to visual mental imagery triggered by natural sounds; however, a similar pattern of results was identified in congenitally blind participants speaking against visual mental imagery explanation (Vetter et al., 2020). Yet, only in the congenitally blind group - but not in the sighted group - the patterns identified in early visual cortices could be generalized across participants (Musz et al., 2021), pointing at differences in employment of these areas by the two groups. Overall, the line of research investigating auditory-induced responses in early visual areas clearly demonstrates a possibility of the signal transfer across the sensory systems.

A different approach could be used to test more specifically the mechanisms enabling activation of representations across sensory modalities. Namely, the representations in two sensory systems need to converge to a high-level abstract representation which can in turn elicit item-specific responses in any of these systems (e.g., the same representation elicited in response to a spoken word “cat” and an image of a cat). Therefore, if we identify where and when modality-independent representations emerge, we will characterize how visual representations are activated by signals belonging to the other modality.

We used a paradigm from object recognition research that could be suitable to study amodal representations. Previously used to track the spatiotemporal dynamics of visual signals (Shinkareva et al., 2011; Fairhall & Caramazza, 2013; Cichy et al., 2014; Kumar et al., 2017; Leonardelli, Fait & Fairhall, 2019), this paradigm can be also utilized to track perception time course in the other sensory systems. Notably, this paradigm allows to test for modality-independent representations, that is, when activity patterns elicited in response to two objects in one modality (e.g., images of two animals) resemble the activity patterns elicited by these objects in the other modality (e.g., hearing their names).

Another close line of research demonstrated a possibility to generalize between natural sounds and natural scenes (Simanova et al., 2012; Jung, Fait & Walther, 2018). While these studies show correspondence in the brain regions supporting cross-modal representations, the inferences provided by these studies are likely to be constrained by the type of chosen stimulation: sounds and views of nature frequently co-occur (Hebb, 1949) throughout human's lifetime, whereas spoken words and natural scenes do not have such a tight audiovisual association and are likely to require other mechanisms for cross-modal activation of representations.

Overall, in order to investigate how visual representations are activated by feedback signals from another sensory system, we conducted Study III (described below in more detail). In this study, we tested a hypothesis that perception signals in two different sensory systems become more conceptual and modality-independent during object recognition, so that there is a point in time when the representations emerging in each modality can be generalized across these modalities.

Altogether, in this work, we investigated how feedback signals coming from within the visual system or from a different sensory system - elicit visual representations. In this regard we addressed two respective research questions: **(1) How are visual feedback representations not confused with visual feedforward input? (2) How are visual representations activated by feedback signals from another sensory system?** Building on the research interrogating the first question, we conducted Studies I and II focusing on feedforward and feedback signal differentiation in cortical depth of V1 using 7T fMRI. Following the research which addresses the second question we conducted Study III, where we explored the mechanism of cross-modal generalization using EEG.

In Study I, we explored differentiation of feedforward and feedback signals by depth in V1 grey matter. We used a mental rotation paradigm where participants were presented with an

oriented grating and were instructed to mentally rotate it in a certain direction for a certain amount of degrees. During the interval given to perform mental rotation, we read out the representations of the perceived (feedforward) and the mentally rotated (feedback) gratings from area V1 in superficial, middle and deep cortical depth bins approximating cortical laminae. Crucially, we used the same trials for assessing both feedforward and feedback signals, thereby avoiding the challenge of isolated signal estimation. We show a clear differentiation of the signal type by cortical depth which corresponds to the signal distribution in cortical laminae in primates: feedforward contents were more strongly represented in the middle compartment of grey matter, whereas feedback contents were more strongly represented in the outer (superficial and deep) cortical bins. Such spatial separation indicates that feedforward and feedback information streams majorly bypass each other within the same brain area which could be a mechanism for precluding a confusion between signals coming from the external world and the internally-generated contents.

Study II built on the results from the Study I to more precisely estimate the signal-by-depth differentiation. In Study II we benchmarked three MR-sequences - gradient-echo, spin-echo, and vascular space occupancy - at 7T fMRI. We used a perceptual flicker paradigm to estimate information mainly carried by feedforward signals and working memory paradigm to estimate feedback signals during working memory maintenance. We demonstrated that the three tested MR-sequences show correspondence in the separation of feedforward/feedback signals by cortical depth: feedforward and feedback information emerged in the middle and deep cortical compartments, respectively. The sharpest differentiation was achieved with the spin-echo sequence based on inter-compartment differences achieved when estimating each of the information streams. We conclude that the spin-echo method might offer a good compromise of

signal sensitivity and spatial specificity for future studies of feedforward/feedback mechanisms at 7T fMRI.

Study III addressed the question of how visual representations are activated by signals belonging to another modality. Specifically, we hypothesized that the neural representations in two different sensory modalities can converge towards conceptual representations that are modality-independent. In Study III, we presented participants with images of objects and spoken words corresponding to the same objects while they were doing an orthogonal 1-back task. Using EEG and time-resolved multivariate pattern analysis we tracked visual and auditory category information to identify objects 'category representations in each modality and across modalities. First, we found robust object category information in both visual and auditory modalities as well as a similar representational transition from individual-object level to a representation of the objects 'category membership in each modality. These results suggest an analogous hierarchy of information processing across sensory channels. Despite the robust representations of objects and object categories in visual and auditory modalities, we did not find evidence for a shared supra-modal code at the level of patterns extracted from EEG scalp electrodes. This result suggests that the contents of the different sensory hierarchies could be ultimately modality-unique.

Overall, the results of our studies show that feedback signals within the visual system propagate all the way down to V1 enabling precise mental transformations (during mental rotation and working memory maintenance in Studies I and II, respectively) via effective delineation of feedback signals from sensory inputs and ultimately initiating distinct subjective experiences. This differentiation can be revealed with diverse MR-sequences at 7T fMRI, where spin-echo sequence could be particularly suitable for establishing cortical depth-specific effects in humans. However, we did not find modality-independent representations which, according to our hypothesis, may

subserve the activation of visual representations by the signals from another sensory system. This pattern of results indicates that identifying the mechanisms bridging different sensory systems is more challenging than exploring within-modality signal circuitry and this challenge requires further studies.

A Chapter 2 “Study I: Perceived and mentally rotated contents are differentially represented in cortical depth of V1” (pages 15-35) was removed for copyright reasons. The text can be found here: https://doi.org/10.1038/s42003-021-02582-4

3 Study II: Benchmarking GE-EPI, SE-EPI and SS-SI-VASO sequences for depth-dependent separation of feedforward and feedback signals in high-field MRI

3.1 Introduction

Recent advances in high-field fMRI have allowed neuroscientists to differentiate feedforward and feedback signals across cortical depth in the healthy human brain. Feedback signals were identified in superficial and/or deep grey matter compartments, whereas feedforward signals were found at the middle cortical depth or across all the cortical compartments (Bergmann et al., 2019; Lawrence et al., 2018; Iamshchinina et al., 2021b). Such depth-specific separation of feedforward and feedback signals has been most firmly established in primary visual cortex (Muckli et al., 2015; Kok et al., 2016; Lawrence et al., 2018; Bergmann et al., 2019; Aitkens et al., 2021; Iamshchinina et al., 2021b), but has also been found in somatosensory cortex (Yu et al., 2019). In most of these studies, the depth-specific signal differentiation was established using gradient-echo echo-planar imaging sequence (GE-EPI). While GE-EPI yields strong blood-oxygen level dependent (BOLD) responses and thus offers high signal sensitivity, it is argued to offer relatively low spatial specificity (Menon et al., 1995; Turner, 2002; Koopmans et al., 2010; Polimeni et al., 2010; Gagnon et al., 2016). As spatial specificity is essential for the precise estimation of depth-dependent cortical profiles, alternative acquisition methods that offer higher spatial specificity have recently attracted attention: particularly, spin-echo (SE)-EPI BOLD (Yacoub et al., 2003; Duong et al., 2003; Olman et al., 2010; Boyacioglu et al., 2014) and the cerebral blood volume (CBV)-based laminar fMRI using the SS-SI-vascular space occupancy (VASO) method (Huber et al., 2014). On the flipside, however, these methods yield reduced signal sensitivity compared to GE-EPI (Zhao et. al., 2006; Huber et al., 2017). As the problem of signal differentiation in grey matter depth requires both high spatial specificity and signal sensitivity, it

is unclear how the tradeoff offered by each of the imaging methods pans out in a typical cognitive neuroscience experiment targeting feedforward and feedback information readout at 7T fMRI.

Here, we benchmarked three MR-sequences at 7T - GE-EPI, SE-EPI and SS-SI-VASO - in their ability to differentiate feedforward and feedback signals in human early visual cortex (V1). The experiment consisted of four participants completing multiple scanning sessions with two complementary tasks: a perception task that predominantly probed cortical feedforward signals, and a working memory task that engages feedback signals. In the perception task, participants viewed oriented gratings while reacting to color changes of fixation cross. In the working memory task, they had to maintain specific grating orientations in working memory. We used multivariate pattern analysis to read out the perceived (feedforward) and memorized (feedback) orientations from neuronal signals across cortical depth. Basing our hypotheses on previous layer-specific work, we expected the perception signals to be predominantly encoded in the middle cortical bin and the working memory signals to emerge in one or both outer cortical bins (superficial and deep). We estimated these effects in every MR-sequence. For comparability of the results across the three imaging methods, we performed the same post-processing steps on the acquired data and estimated the effects on closely matching time intervals in the trial.

Data analysis across all of the MR-sequences revealed the feedforward signal predominantly at the middle cortical compartment of area V1 and the feedback signal in the deep compartment. The results obtained with conventional GE-EPI generally agree with the results obtained with VASO and SE-EPI, indicating that GE-EPI is a suitable method for the exploration of signals in cortical depth with 7T fMRI. The VASO sequence showed weak representations of both feedforward and feedback signals, although the overall pattern of results corresponded to those found for the other sequences. Interestingly, SE-EPI was the only sequence yielding

statistically reliable differences between the cortical compartments that show feedforward and feedback modulations. We therefore suggest that in the context of a typical cognitive experiment such as here, SE-EPI may provide a suitable trade-off between spatial specificity and signal sensitivity.

3.2 Methods

3.2.1 Participants

Four adults (age in range from 28 to 35; 2 female) participated in the study. All participants had normal or corrected-to-normal vision. Each participant had extensive experience with participating in psychophysics studies at high-field fMRI and participated in two sessions per MR-sequence (GE, SE, VASO), resulting in 6 sessions per participant and 24 sessions in total. Participants gave their written informed consent for participation in the study as well as for publicly sharing all obtained data in pseudonymized form. They received monetary reimbursement for their participation. One of the participants did not complete all the perception runs during the first session of GE and first session of SE measurements, and thus only the data from the second sessions of both sequences is utilized for the feedforward signal estimation. The study was approved by the ethics committee of the University of Leipzig, Germany.

3.2.2 Stimuli

Stimuli were grayscale luminance-defined sinusoidal gratings generated using MATLAB (MathWorks, Natick, MA) in conjunction with the Psychophysics Toolbox (Brainard, 1997). The gratings were presented in an annulus (outer diameter: 6.7° of visual angle, inner diameter: 1.3° of visual angle) surrounding a central fixation point. The gratings had a spatial frequency of 2 cpd (12.34 Hz) and a Michelson contrast of 50%. Stimuli were displayed on an LCD projector (DLR-

RS49E, JVC Ltd.) on a rear-projection screen positioned behind the head coil within the magnet bore. Participants viewed the screen through a mirror attached to the head coil.

3.2.3 Experimental procedure

3.2.3.1 Training procedure

Before entering the MRI scanner, participants underwent a training procedure which comprised a minimum of 4 runs for all the participants. At the start of each trial, participants briefly saw two randomly oriented gratings (Figure 1). A subsequently presented task cue indicated which of the gratings needed to be memorized for the upcoming task. After an 11 second retention time period a probe grating was shown. The grating was slightly tilted clockwise or counterclockwise with respect to the grating that had been cued; the amount of additional tilt was regulated in a staircase procedure (described below). The participants' task was to indicate whether the probe grating was tilted clockwise or counterclockwise from the memorized grating. After each trial, participants received a 1 second feedback about their response correctness. The inter-trial interval was 2 seconds. Each training run consisted of 16 trials and took 4 minutes 54 seconds. At the end of each run, participants received feedback about their average accuracy.

3.2.3.2 Experimental perceptual task

To specifically investigate visual processing dominated by feedforward signals, we included two perceptual runs in the experiment. During these runs, gratings with the two target orientations (25° and 115°) were shown in a pseudo-randomized order. On each trial, one of the grating orientations was shown for 15 seconds, flickering at 2 Hz. The presentation was followed by a fixation cross, which lasted 9 seconds in GE and SE measurements and 10 seconds in VASO measurements. Participants had to monitor the fixation cross for occasional brief changes in color, to which they had to respond with a button press. Overall, on every run we recorded 16 trials

evenly split between the two orientations conditions. The perceptual task was performed in the experiment during runs 3 and 6 (out of 9 runs overall). The fixation dot changed 7.4±0.03 times per trial at random time points, leading to approximately 118 changes, to which participants responded on average 75+/-3% (Mean ± SD) of the time. As for the experimental runs, each perceptual task run took 6 minutes 42 seconds for GE and SE measurements and 7 minutes for the VASO measurements.

3.2.3.3 Experimental working memory task

In the scanner, participants continued to perform the working memory task they were trained on (Figure 1), but with three major changes. First, participants did not receive feedback on their performance. Second, the ITI was prolonged to 9 seconds for the measurements with GE and SE sequences and to 10 seconds with VASO sequences. This was done to reduce the temporal overlap between the perceptual trace from the probe grating shown on the previous trial and the current trial. The length of each trial (including ITI) was 24 seconds in the GE and SE sessions and 25 seconds for the VASO sessions. The difference in trial lengths also ensured that every trial was evenly divisible by a fixed number of TRs: with a 3 second TR used for GE and SE measurements, each trial comprised exactly 8 TRs; whereas with 5 seconds TR used for the data acquisition with VASO, each trial comprised exactly 5 TRs. Third, the sample gratings shown at the beginning of each trial were no longer randomly oriented, but always either 25° or 115° orientated away from the vertical axis. We limited the number of possible orientations compared to the training session to increase the signal-to-noise ratio for each grating orientation and thereby enable us to differentiate orientation signals at the level of cortical depth bins.

Every run consisted of 16 trials. In each half of trials, one of the two possible orientations (25° versus 115°) was selected for memorization. Trial order was fully randomized. The

experiment consisted of 7 runs, which each lasted 6 minutes 42 seconds for GE and SE and 7 minutes for VASO. The average time for completing the whole experiment in one session was 80-85 minutes.

The anatomical scans for each participant were acquired separately during other studies in MPI Leipzig. None of these scans were older than 6 months.

We invited every participant two times per scanning sequence (GE, SE and VASO) to increase signal-to-noise ratio. Each participant was thus scanned in 6 experimental sessions. The order of the scanning sequences was randomized across the sessions for every subject.

3.2.3.4 Staircase Procedure

To maintain a sensitive accuracy range across the whole experiment, including the training runs and the fMRI experiment, we used a staircase procedure that adjusted the amount of additional tilt in the probe grating compared to the memorized grating. The initial difference between the memorized orientation and the probe grating was set to 20°. For each correct response in each subsequent trial, the difference between the probe and rotated grating was reduced by 0.5°, making orientation discrimination harder. Conversely, the difference was increased by 2° for each incorrect response, making discrimination easier. We imposed an upper limit of 40° on the orientation difference. Across all the MR-sequences participants' performance ranged in between (mean ± standard deviation): 8.8°±5.1° .

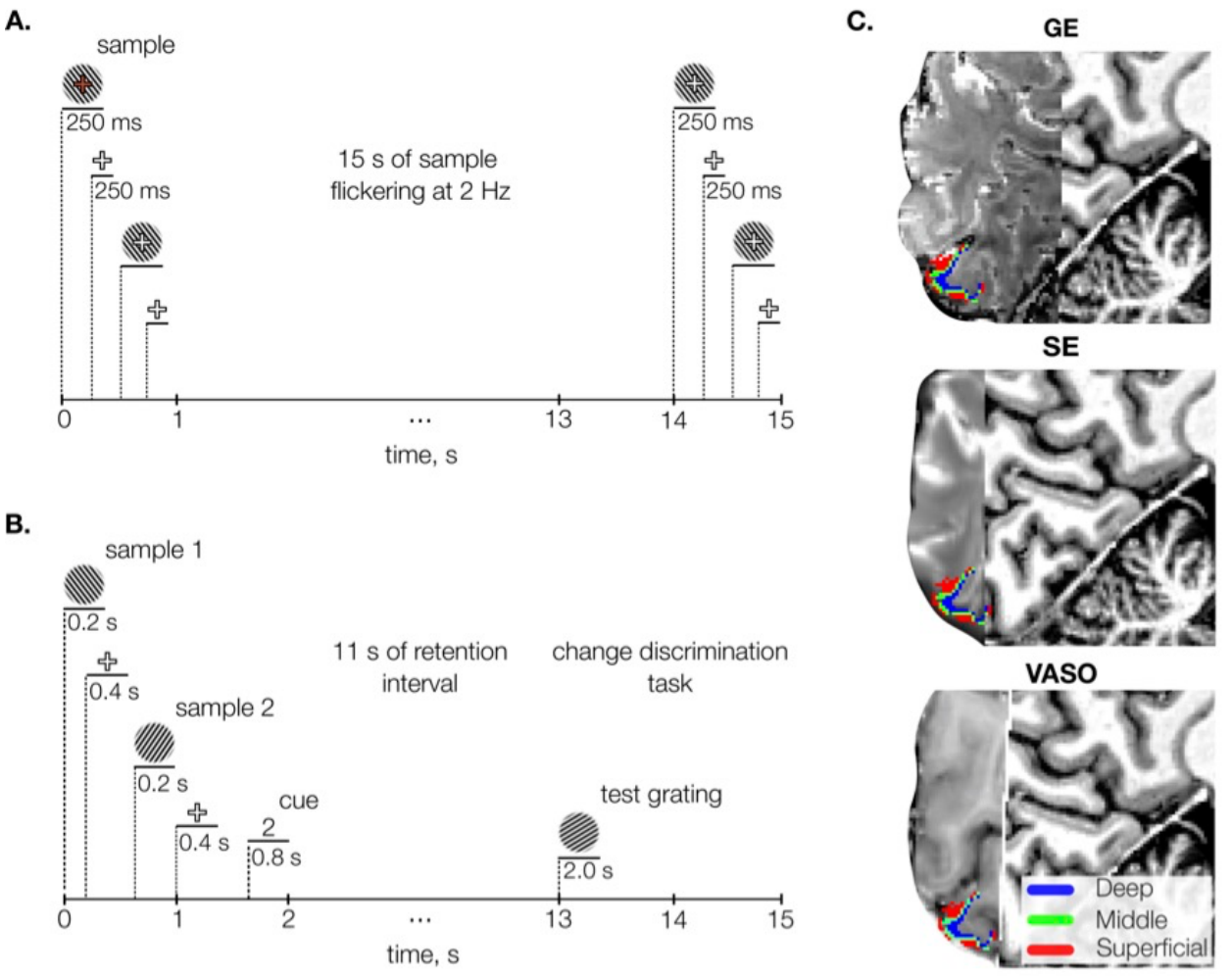


Figure 1. Methods. A. Perceptual task. On each trial participants viewed a sample grating flickering at 2 Hz. They had to press a button when the fixation cross was turning red. **B. Working memory task.** On each trial, participants viewed two sample gratings, then a cue ("1" or "2" in the screen center) indicated which of the grating orientations was to be memorized. After a retention interval of 11 seconds, a probe grating was shown, and participants had 2 seconds to report whether the probe was tilted clockwise or counterclockwise compared to the memorized grating. **C.** Sagittal slice of the anatomical image of a representative participant overlaid with the three average functional images acquired with different MR-sequences (GE-EPI, SE-EPI, VASO) and cortical depth bins approximating cortical layers (superficial, middle and deep) from an equi-volumetric model (see Methods). The cortex is mapped within the region of V1 with voxel eccentricity values 1-3°.

3.2.4 Parameters of data acquisition

fMRI was acquired on a 7-T Magnetom (Siemens Healthineers, Germany) whole-body scanner using a single-channel-transmit/32- channel RF receive head coil (Nova Medical Inc, USA). fMRI data were recorded with an isotropic spatial resolution of 0.8 mm using GE-EPI, SE-EPI and SS-SI-VASO. For the GE-EPI protocol we used the CMRR MB sequence (TR 3000 ms, TE 24 ms, number of slices 50, 78/90° flip angle, 148 X 148 mm^2 FOV, GRAPPA acceleration factor 3, slice partial Fourier 6/8, coronal orientation, F >> H phase encoding direction) (Moeller et al., 2010; Feinberg et al., 2010). For the SE-EPI protocol we used the following parameters (TR 3000 ms, TE 38 ms, number of slices ~30, 90° flip angle, 148 X 148 mm^2 FOV, GRAPPA acceleration factor 3, slice partial Fourier 6/8, coronal orientation, F >> H phase encoding direction). For the SS-SI-VASO protocol we used the following parameters (TR 2837.90 ms, TE 25 ms, TI 650 ms, number of slices 26, 26° flip angle, 133.0 X 133.0 mm^2 FOV, GRAPPA acceleration factor 3, slice partial Fourier 6/8, orientation T > C-28.2, A >> P phase encoding direction). Shimming was performed using the standard Siemens procedure. Anatomical data were acquired using a MP2RAGE sequence with 0.7 mm isotropic resolution (TR 5000 ms, TE 2.45 ms, TI 900/2750 ms, flip angle 5/3, bandwidth 250 Hz/Px, 224 × 224 mm FOV, GRAPPA acceleration factor 3, slice partial Fourier 6/8, base resolution 320, sagittal orientation, A>>P phase encoding direction, scan time 10 min 57 sec).

3.2.5 Functional and anatomical data preprocessing

3.2.5.1 Bias field correction and segmentation of the anatomical image

The DICOM data were converted to NIfTI format using SPM12 (Wellcome Trust Center for Neuroimaging, University College London). The volumes were bias field-corrected using a SPM-based customized script (Luesebrink, Sciarra, Mattern 2017). To implement cortical depth-

specific analysis, we extracted grey matter segmentation for each subject. To do this, first we used the SPM 12 segmentation algorithm and then the brainmask was generated by adding up the white matter, grey matter and cerebro-spinal fluid masks. Then we applied the FreeSurfer (version 6.0.0) recon algorithm to perform segmentation of white matter, grey matter, generating their surfaces and a binary brain mask of the cortical ribbon (1 if the voxel falls into the ribbon, 0 otherwise (steps 5-31 of recon-all algorithm)). We ran the recon algorithm on the extracted brainmask from a T1-weighted image with a '-hires' flag for the data with resolution higher than 1 mm (Zaretskaya et al., 2017). After running the recon algorithm, the Freesurfer-generated grey and white matter segmentations were visually inspected in each participant, the borders between CSF and grey matter as well as grey matter and white matter were manually corrected within the region corresponding to the field of view of functional scans.

3.2.5.2 Cortical depth and ROI definition

The grey matter segmentation acquired with Freesurfer was further utilized to obtain cortical depth-specific compartments. Deep, middle and superficial compartments were constructed using an equi-volumetric model (Waehnert et al., 2014; Huntenburg et al., 2018). In order to analyze depth-specific activity in early visual areas, we applied a probabilistic surface-based anatomical atlas (Benson et al., 2014) to reconstruct the surfaces of area V1 for each subject. This is an atlas of the visual field representation (eccentricity and polar angle), and eccentricity values were used to select the foveal sub-part of the surface excluding the area occupied with the fixation cross (1-3°). The extracted surface ROI (V1) was then projected into the volume space and intersected with the predefined cortical compartments. In this way, we obtained the V1 ROI in the Freesurfer anatomical space at three predefined cortical depths.

3.2.5.3 Functional data preprocessing

Functional volumes were realigned to the first volume of the middle run using SPM software (Statistical Parametric Mapping; SPM12). In the data acquired with VASO sequence, the nulled and not nulled frames were realigned separately (Finn et al., 2019), the not-nulled volumes were interpolated using 7th order spline function to correspond to the acquisition time of the nulled frames (the nulled volumes did not undergo temporal interpolation in our study, in order to reduce the sharing of informational content across time points). Motion traces of nulled and not-nulled frames were then visually inspected to ensure good overlap of these contrasts. The nulled data was then corrected with the not-nulled volumes using the dynamic division method (Finn et al., 2019).

After the realignment, we ensured that the functional runs acquired with each scanning sequence were well aligned with each other in each participant. This is required for multivariate pattern analyses of high-resolution fMRI data. For this we computed inter-run spatial cross-correlations of the signal intensities of the functional volumes. The resulting average spatial correlation of experimental runs was very high: (Mean ± SD) 0.97±0.03 in GE-EPI sessions, 0.986±0.01 in SE-EPI sessions and 0.99±0.004 in SS-SI-VASO sessions. Further, functional-anatomical alignments were checked visually to ensure that the functional scans were well aligned to the anatomical image at the location of the ROIs.

After that, the data acquired with all the three sequences was high-pass filtered (removing signal with $f<1/128$ Hz). Before classification analysis, the functional time series of every voxel within the ROI was z scored to correct for the scaling differences in voxel intensities within every run (Lawrence et al., 2018).

3.2.5.4 Registration

We linearly coregistered the extracted ROIs with predefined cortical depth compartments to the EPI volumes within each subject using the Symmetric Normalization (SyN) algorithm of ANTs (Avants et al., 2008). Specifically, first, the T1-weighted anatomical image was registered using linear interpolation to the EPI volume averaged over all the functional runs. Next, we registered the ROIs with the predefined cortical depths to the EPI volume using nearest neighbor interpolation and by applying the coordinate mapping (with the voxel size resampled to the functional runs (0.8 isotropic)) obtained in the previous step. The resulting ROIs included the following number of voxels per cortical depth and MR-sequence (Mean ± SD): GE-EPI (M_{deep}=1170.8±90, M_{mid}=1083±62, M_{super}=1170.2±109), SE-EPI (M_{deep}=1026.1±197, M_{mid}=955.1±158, M_{super}=1035±216), SS-SI-VASO (M_{deep}=1120.8±97, M_{mid}=1037±95, M_{super}=1120±143).

3.2.6 Multivariate pattern analysis

3.2.6.1 Data extraction

Multivariate pattern analysis (MVPA) was performed in each subject and experimental session separately. To prepare the EPI data for the MVPA, we first extracted activity patterns for a V1 ROI with the predefined cortical depths from the functional images in the experimental task runs. Specifically, in each run, we extracted voxel-wise activation values for the 2 oriented grating conditions (25° and 115°) and 8 trials for each condition across trial TRs starting at trial onset.

3.2.6.2 Classification

MVPA was carried out using linear support vector machines (SVMs; libsvm: http://www.csie.ntu.edu. tw/~cjlin/libsvm/) with a fixed cost parameter (c=1). We performed classification both across all the voxels in the full grey matter ribbon of V1 (Supplementary Figure

1) as well as separately at each cortical depth of V1. To this end, we trained the SVM classifiers on multi-voxel response patterns from all the experimental runs within each experimental session, leaving out one run (i.e., using leave-one-out cross validation), to discriminate between the 2 oriented gratings for each time point in the trial. All trials in the training set were utilized for the classifier training (8 trials per orientation and training set in perception task and 48 trials per orientation and training set in working memory runs). Next, we tested the SVM classifier using the trials of the left-out run (8 trials in both tasks). The classifier was trained on each time point in the trial using the data from the training set and tested on the corresponding time point in the test set. As a result, we extracted decoding accuracy for every TR of all the runs in the main experiment (chance level 50%). The results were organized in the form of the matrix for further statistical testing: participants (4) x experimental runs (number of runs depended on the working memory (7) or perception task (2)) x TRs (number of TRs depended on the scanning sequence (GE - 8; SE - 8; VASO - 5)).

3.2.6.3 Statistical testing

For the statistical assessment of feedforward and feedback effects, we preselected the time interval in the trial of a respective task. We estimated the feedforward effect by averaging over the classification accuracy obtained during the 12 seconds interval following stimulus presentation in the perception task trials. In GE- and SE-acquired data, the measurements at 6, 9 and 12 seconds were included in this analysis (Supplementary Figure 1). The first time point in the trial (after 3 seconds) was excluded as uninformative, as it was too close to the stimulus presentation to carry reliable orientation information. In the VASO data, the classification accuracies obtained at 5 and 10 seconds in the trial were averaged and underwent further statistical testing. Here, we did not

exclude the first time point (after 5s), since this later time point was likely to already contain information about the stimulus presentation.

For the statistical assessment of feedback contents, we preselected the critical time interval based on previous studies (Harrison & Tong, 2009; Albers et al., 2013; Rademaker, Chunkaras & Serences, 2019) where working memory representations could be decoded starting from 6 seconds after stimulus onset. For GE and SE measurements, the critical interval in our study included measurements at 6, 9, 12 and 15 seconds, that is, all the time points starting from 6 seconds until one time point after the probe grating onset (at 13 seconds) (Supplementary Figure 1). We also included the time point after the probe grating presentation, since this measurement was too close to the presentation of the probe grating to be contaminated by it (Iamshchinina et al., 2021b). For the statistical analysis of VASO-based estimates, we included the time points starting from 6 seconds or later, which resulted in measurements at 10 and 15 seconds.

To test whether decoding of orientation in feedforward and feedback signals was significantly above chance and to compare the decoding between cortical depth bins, we used a linear mixed modeling approach (fitlme function, Matlab, The Mathworks Inc, 2014). First, in order to obtain a baseline against which the presence of the signal could be tested, we generated permuted null distributions of classification accuracies for each participant, timepoint and sequence. On each permutation, we reshuffled orientation labels for all trials before performing the classification (1,000 iterations with shuffled data labels). Then, we averaged the classification results of all the permutations and compared them to the original classification results (the empirical effect). The feedforward and feedback conditions were estimated separately due to substantial differences in the paradigms and trial time intervals chosen. The data from all the runs of all the participants were concatenated within each experimental session without averaging (i.e.,

16 data points in the perception task (4 subjects X 2 sessions per imaging method X 2 experimental runs) and 56 data points in the working memory task (4 subjects X 2 sessions per imaging method X 7 experimental runs)). The linear mixed effects model included an intercept, classification accuracy as a response variable, cortical depth variable as a fixed effects portion of the model (categorical predictor), participants number as a grouping variable with the random-effects terms for intercept and cortical depth specified separately (assuming no correlation between them): classification accuracy ~1+depth_effect + (1|participants) + (depth_effect-1|participants).

3.3 Results

In this study, we compared the performance of three MR-sequences (GE, SE and VASO) in distinguishing feedforward and feedback signals in cortical depth of area V1. Using each MR-sequence and 7T fMRI, we measured feedforward and feedback-dominated information emerging during perception and memorization of orientated gratings, respectively (Figure 1). We collected data from 4 participants, where each participant undertook two sessions per scanning method. Signal estimation was performed in separate grey matter compartments (superficial, middle and deep) extracted with an equi-volumetric model (see Methods). For each depth bin, we trained support vector machine classifiers to differentiate multi-voxel response patterns evoked by the different grating orientations. For this analysis, we focused on time points which were expected to carry robust perceptual signals (measurements at 5 and 10 seconds in VASO measurements and for other sequences: 6, 9 and 12 seconds) or memory signals (measurements at 10 and 15 seconds in VASO measurements and for other sequences: 6, 9, 12 and 15 seconds). By comparing classification in these time windows between the grey matter bins, we establish which of the MR-sequences uncovers differences in feedforward and feedback signals across cortical depth.

3.3.1 Feedforward signals across cortical depth in V1

We examined the cortical profile of orientation-selective activity measured during the perception task with the three MR-sequences. Firstly, we established whether the different perceived orientations could be successfully decoded at different cortical depth. Secondly, following the predictions based on previous animal and human work (Rockland & Pandya, 1979; Lawrence et al., 2019), we hypothesized that the middle cortical bin represents the perceived contents more strongly than the outer cortical bins averaged together.

Analysis of the GE-EPI data yielded above-chance classification results in the middle cortical bin ($t(26)=2.6$, $p=0.01$) but not in the other bins ($p>0.1$). There was also no significant difference between the cortical bins ($p>0.4$) (Figure 1A).

Analysis of the SE-EPI data revealed a highly significant above-chance classification results in the middle ($t(26)=5.0$, $p<0.001$) and superficial bins ($t(26)=4.2$, $p<0.001$) but not in the deep bin ($p=0.7$). In accordance with our hypothesis about a differential representation of the feedforward signal, we observed a higher classification accuracy for perceived orientation in the middle bin than in the superficial and deep bins averaged together ($t(26)=2.7$, $p=0.01$). Unpacking this effect, we found a higher classification accuracy for perceived orientation in the middle bin than in the deep bin ($t(26)=4.9$, $p<0.001$), with no difference between the middle and the superficial bins ($p=0.7$).

Finally, analysis of the VASO data yielded above-chance classification results for perceived orientation in the middle cortical bin (on the trend-level; $t(30)=2.0$ $p=0.06$) and not in the other bins ($p>0.2$), with no significant differences between the cortical bins ($p>0.3$).

Overall, we localized feedforward signals at the middle cortical depth bin of V1 for all the MR-sequences. However, we only observed a significant feedforward representation in the middle

cortical compartment of V1 when the data were acquired with SE-EPI, but not with the other sequences. This differential feedforward representation for SE-EPI was mainly driven by a difference between the middle and deep cortical bins.

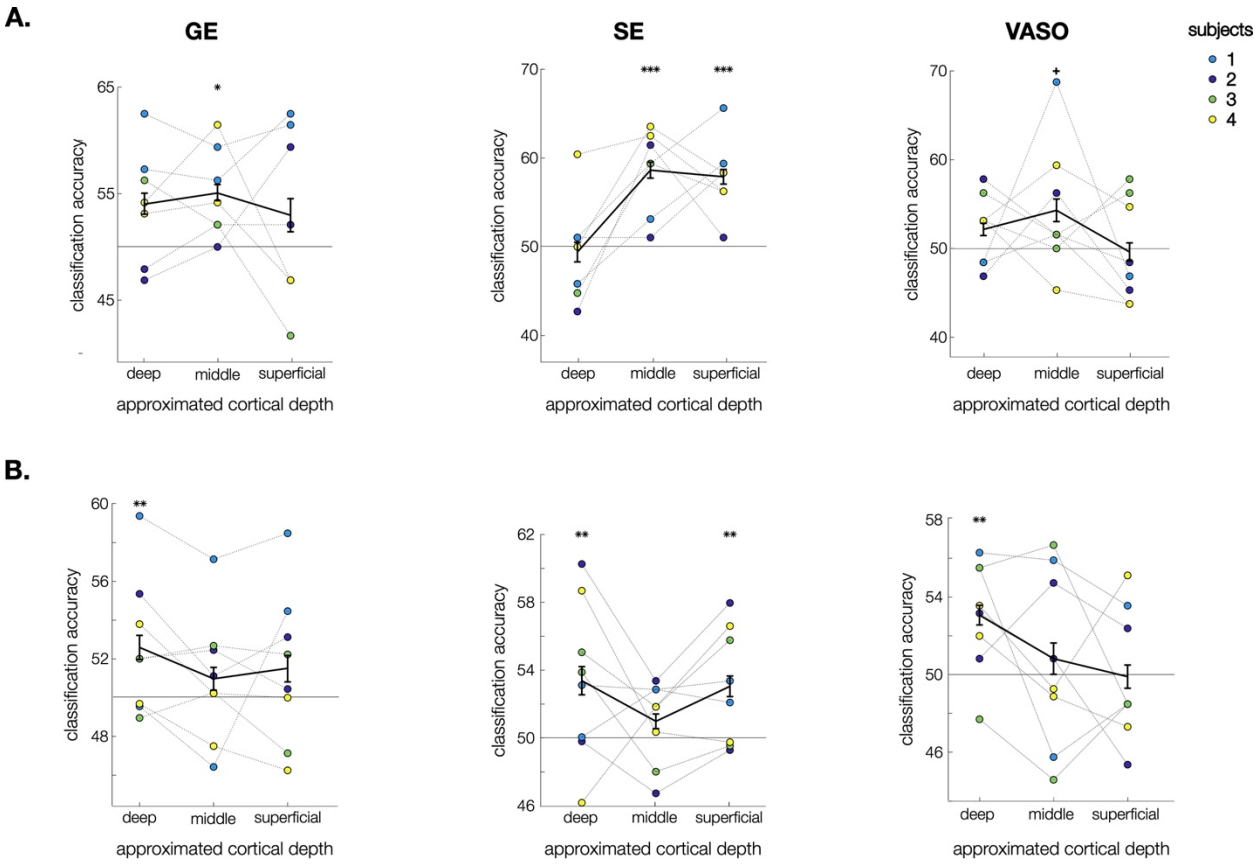


Figure 2. Results. A. Classification accuracy for decoding of perceived grating orientation in V1 measured with GE-EPI, SE-EPI and VASO. Above-chance classification results for perceived orientation were found in the middle cortical bin in all the sequences and at superficial depth for SE-EPI. Significant differentiation across cortical depth only emerged for SE-EPI, where decoding was stronger in the middle bin than in the other two bin combined. **B.** Classification accuracy for decoding of memorized grating orientation. Above-chance classification results for memorized orientation were found in the deep cortical bin in all the sequences and in the deep and superficial cortical bins for SE-EPI. A significant differentiation between the middle and outer depth bins was again only achieved using SE-EPI. The colored dots represent experimental sessions (2) for each participant (4). The data points are to show the correspondence between two separate sessions within each participant. Note that the statistical analysis was performed using experimental runs unaveraged within each of the sessions. All error bars denote standard error of mean within

participants. Represented p-values are fixed effects coefficients estimated with a t-test for each cortical depth (uncorrected) +: p<0.07, *: p<0.05, **: p < 0.01, ***: p < 0.001.

3.3.2 Feedback signals across cortical depth in V1

Next, we examined differences in revealing feedback representations in human V1 across the three scanning sequences. We analyzed the orientation-selective activity acquired during the working memory task focusing our analysis on the retention time interval (see Methods). As for the analysis of the perception signal, we aimed to establish which cortical compartments carry the representation of the memorized item. Firstly, we established whether the different memorized orientations could be successfully decoded at different cortical depth. Secondly, following our predictions about the layer-specific distribution of feedback signals, we tested if any of the outer cortical bins represented the memorized contents more strongly than the middle cortical bin.

Analysis of the GE-EPI data revealed above-chance classification results in the deep cortical bin ($t(104)=2.8$, $p=0.007$) but not in the other bins ($p>0.2$), with no significant difference between the bins ($p>0.3$).

Analysis of the SE-EPI data yielded above-chance classification in both the deep ($t(102)=2.7$, $p=0.008$) and superficial ($t(102)=2.9$, $p=0.004$) bins, but not in the middle bin ($p>0.26$). Comparing feedback representations across cortical depth revealed that classification accuracy for memorized orientation in the outer bins averaged together was higher than the in middle bin (at the trend level; $t(102)=1.8$, $p=0.07$). Pairwise comparison between all three cortical compartments showed higher classification accuracy for memorized orientation in the superficial bin than in the middle bin at the level of trend ($t(102)=1.9$, $p<0.06$) with no difference between the rest of the bins ($p>0.15$).

Finally, analysis of the VASO data revealed above-chance classification results for memorized orientation only in the deep cortical bin ($t(110) = 2.6$, $p = 0.009$) but not in the other cortical bins ($p>0.5$), with no difference between the cortical compartments ($p>0.6$).

Overall, we observed feedback signals in the deep cortical compartment of V1 for all the MR-sequences. However, we only observed a differential representation of feedback signals across the middle and outer cortical compartments of V1 when the data was acquired with SE-EPI, but not with the other sequences. This differential feedback representation for SE-EPI was mainly driven by a difference between the superficial and middle cortical bins.

3.4 Discussion

In the present study, we investigated depth-dependent separation of feedforward and feedback signals at 7T fMRI using three MR-sequences: GE-EPI, SE-EPI and SS-SI-VASO. We used multivariate pattern classification to read out information about the contents of feedforward and feedback information across cortical depth in area V1.

For all three sequences, we were able to decode feedforward signals from the middle depth bin of area V1, while feedback could be read out from the deep cortical bin. This consistency indicates that the widely used GE-EPI (Muckli et al., 2015; Kok et al., 2016; Lawrence et al., 2018; Bergmann et al., 2019; Aitkens et al., 2021; Iamshchinina et al., 2021b) is a viable method for decoding information from cortical depth bins. This suggests that despite its lower spatial specificity than SE-EPI or VASO (Yacoub et al., 2003; Duong et al., 2003; Olman et al., 2010; Huber et al., 2014), the more sensitive GE-EPI measurements can detect depth-specific signals in V1.

From the three sequences SE-EPI stood out by being the only sequence in our study that yielded a statistically reliable depth-specific separation of the feedforward and feedback signals,

where feedforward signals were significantly stronger in the middle bin compared to the outer bins and feedback signals were stronger in the outer bins compared to the middle bin. These results align well with previous animal research and 7T work in humans (Kok et al., 2016; Aitken et al., 2021; Bergmann et al., 2019) and highlight SE-EPI as a particularly interesting MR-sequence for establishing layer-specific effects in experiments of the kind used here, through its favorable trade-off between sensitivity and spatial specificity.

The depth differentiation was mainly reflected in differences between the middle and deep bins. The estimates obtained in the superficial depth bin were found to be less consistent across the MR-sequences. This could be related to a bias towards the surface veins observed in the SE-EPI measurements (Supplementary Figure 3, Uludag, Mueller-Bierl, Ugurbil, 2009) which could have impacted the effects obtained at the superficial cortical depth. Alternatively, however, the generally lower classification obtained with VASO may have missed signals at the superficial depth. Future studies need to establish whether feedback signals at the superficial cortical depth can be disentangled from nonspecific contributions such as surface-vein effects. For this, they could use model-based approaches (Markuerkiaga, Bath & Norris, 2016; Havlicek & Uludag, 2020; Markuardt et al., 2017; Huang et al., 2021) or combine high-field MRI with different imaging methods (Kashyap et al., 2020; Huber et al., 2018).

While SE-EPI yielded the clearest depth separation in the current benchmark study, several other aspects need to be taken into consideration when deciding between imaging protocols for future studies. First, SE-EPI sequences typically yield reduced BOLD sensitivity compared to GE-EPI (Boyacioglu et al., 2014). In this context, it is worth noting that our study employed a design with few participants with long acquisitions each, which may be the experimental regime particularly suited for SE-EPI. Second, they are affected by unspecific factors, such as residual

venous signal contributions, sensitivity to B1-inhomogeneity, which can dramatically affect the quality of acquired data, and are limited by specific absorption rates, which can drastically reduce the number of slices that can be acquired (van der Zwaag et al., 2014; Marques & Norris, 2018). Given these limitations, future studies could employ improved protocols for SE-EPI looking to overcome currently present challenges (Norris et al., 2011; Mugler, 2014; Gagoski et al., 2015; Han et al., 2021).

In our experiment VASO data showed low decoding accuracy for feedforward and feedback signals (Supplementary Figure 2 and 3). This unfavorable outcome could be partially due to lower temporal resolution of VASO measurements and thus lower statistical power in its analysis compared to other MR-sequences. Specifically in our study, the selection of time intervals for the assessment of feedforward and feedback cortical profiles was more challenging for VASO, because each TR covered larger temporal chunks within the trial (see Supplementary Figure 3).

In sum, whether GE-EPI, SE-EPI or VASO are the most appropriate choice for dissociating feedback from feedforward signals will in practice depend on multiple factors of the experimental context. Our study contributes to mapping out the choice space by providing a single point in this space through benchmarking for a typical cognitive neuroscience experiment – however, future investigation is needed to populate this choice space further, with different protocols and paradigms.

A Chapter 4 “Study III: Resolving the time course of visual and auditory object categorization” (pages 56-68) was removed for copyright reasons.

5 GENERAL DISCUSSION

In this work, we aimed to understand how feedforward and feedback mechanisms support complex visual cognition in humans. To this end, we conducted three studies to explore how feedback modulations coming from within the same or from a different sensory system activate visual representations.

Studies I and II explored how the brain distinguishes feedforward and feedback representations within the visual system using different experimental paradigms and methods at 7T fMRI. In **Study I**, we focused on the mechanisms for separation of feedforward and feedback representations in cortical depth of V1 grey matter. We demonstrated that in the mental rotation task, feedback contents were predominant in the outer cortical depth compartments (i.e., superficial and deep), whereas perceived contents carried mainly by feedforward signals were more strongly represented at the middle cortical compartment. This result demonstrates that concurrently represented information streams are separated by the cortical depth within the same brain region, the mechanism which may contribute to the delineation of the sensory input and the internally-generated contents.

In **Study II**, we compared three MR-sequences with varying signal sensitivity and spatial specificity - gradient-echo, spin-echo and vascular space occupancy - in their ability to differentiate feedforward and feedback signals by cortical depth in V1. The resulting distribution of the signals in grey matter depth was consistent across all the three imaging methods: representations of perceived stimuli carried by feedforward signals emerged in the middle cortical compartment of area V1, while feedback signals (representation of an item held in working memory) were read out from the deep cortical compartment. Such correspondence among imaging techniques corroborates that the widely used gradient-echo (Muckli et al., 2015; Kok et al., 2016;

Lawrence et al., 2018; Bergmann et al., 2019; Aitkens et al., 2021; Iamshchinina et al., 2021b) is an appropriate method for decoding information from cortical depth bins. Critically, spin-echo sequence not only allowed to identify the reliable presence of the signals in the corresponding cortical compartments but also yielded reliable differences between the cortical compartments in representing the feedforward and feedback contents. This result indicates that the spin-echo method might offer a trade-off of signal sensitivity and spatial specificity suitable for a typical experiment manipulating feedforward/feedback signals at 7T fMRI.

Study III explored how signals from a sensory modality other than vision elicit responses propagating through the visual system. Specifically, we tested a hypothesis that signals from two sensory systems can obtain modality-independent feature space which could be read out by both systems thereby enabling signal transition across them. Using EEG we tracked the time course of object category information in visual and auditory sensory systems to identify a timestamp of cross-modal generalization. We successfully identified object and object category representations in each modality by tracking auditory and visual signals in the course of object recognition. However, we did not observe significant cross-modal decoding and, thus, we did not find evidence for a transformation of representations from modality-specific codes to modality-independent conceptual representations.

A substantial body of previous research demonstrated that top-down signals activate neural representations within the visual brain regions as a result of mental transformation (Albers et al., 2013; Christophel et al., 2014; Christophel et al., 2018; Rademaker et al., 2019). Crucially, in Studies I and II we demonstrate how V1 functionally segregates external and internally-generated visual contents through cortical depth compartmentalization. Such signal segregation might underlie the differences in subjective experience of perceived and mentally represented contents,

whereas confusion of feedforward and feedback signals might cause symptoms of psychosis including hallucinations and delusions observed in diverse neurological conditions (Sterzer et al., 2018; Haarsma, Kok & Browning, 2018). Further research is needed to establish the contribution of layer-specific signal differentiation in shaping coherent representation of the outer world.

The results of Studies I and II align well with previous research investigating feedforward and feedback signals in V1 (Lawrence et al., 2018; Lawrence et al., 2019; Aitkens et al., 2021; Bergmann et al., 2018; Kok et al., 2016; Muckli et al., 2015; Huber et al., 2017). Notably, these studies utilized diverse paradigms manipulating the type of feedback (e.g., mental imagery, expectations, working memory). The consistency in results indicates that area V1 is not sensitive to the high-level top-down signals carrying information about a specific cognitive process which conjures up the low-level representation. However, previous 3T fMRI studies demonstrated an ability to extract other types of high-level information from area V1, such as categorical divisions (Williams et al., 2007; Vetter et al., 2014; Morgan & Muckli, 2016) and task-driven information (Seidemann & Geisler, 2018). A promising avenue for for future research is to investigate whether specific cognitive operations (e.g., expectation vs. short-term maintenance) can be differentiated at the level of V1 and whether the high-level information about a type of cognitive operation arrives in a different cortical depth compartment compared to the item-specific information, as predicted from neuroanatomical models (Markov et al., 2014; Bergman et al., 2019).

Studies I and II challenge the previously suggested idea that mental images are ‘‘perception-like’’ which was initially based on the observation in 3T fMRI studies that short-term memory and mental imagery contents resemble those evoked by bottom-up visual stimulation (Cichy, Heinzle & Haynes, 2012; Albers et al., 2013; Xu, 2021; Iamshchinina et al., 2021a). These finding in 3T fMRI studies corresponds to the observation in 7T fMRI research (Iamshchinina et

al., 2021b; Lawrence et al., 2018) that perception signals are represented not only in the middle cortical depth but can also be found in the outer cortical compartments dominated by feedback signals. This indicates an overlap in the in-depth distribution of the signals which can impact subjective experiences. However, the results of Studies I and II demonstrate that feedforward/feedback information streams dominate distinct cortical depth bins. Such a system where signals are only partially separated might be functionally beneficial - aiding subjective vividness of internally-generated representations through their partial overlap with the perceived representations, on the one hand, yet allowing for a reliable signal differentiation, on the other. Measuring feedforward/feedback signals using MR-sequence with improved spatial specificity in the future studies may refine our understanding of the crosstalk between feedforward/feedback information streams and its effects on cognitive functioning by gaining sharper signal differentiation and more precise mapping between cortical depth and observers' subjective experiences (Persichetti et al., 2020).

Studies I and II further offer a refined methodology for investigating feedforward/feedback differentiation in the human brain. They provide both the paradigm which effectively separates these signals in cortical depth and the MR-sequence which is conceivably more suitable for this purpose. Future studies could employ this combined methodology to establish a connection between feedforward/feedback signal transmission in cortical depth and across frequency-bands in humans. Figure-ground segregation studies in primates show that directed gamma-locked influences first emerge in the middle layer of area V1 and directed alpha-locked influences from V2-V4 first enter V1 through superficial and deep layers (Kerkoerle et al., 2014; Bastos et al., 2018) - thus revealing a correspondence between the rhythms and cortical layers supporting them. Such correspondence has not yet been observed in humans. Recently developed paradigms allow

localizing feedforward/feedback signals within specific frequency bands using MEG (Stauch & Fries, 2021). Together, these methodologies would enable a spatiotemporally resolved view on feedforward and feedback dynamics in humans.

In contrast to the results obtained within the visual system, we were not able to uncover how signals from one sensory system trigger selective responses in the other sensory system. Our hypothesis in Study III that auditory and visual representations converge to an abstract modality-independent representation in the course of object recognition, was not confirmed. One possible reason is that object recognition represents a tangled interaction of driving feedforward and weaker modulatory feedback signals (Kar & DiCarlo, 2021). Since it is feedback signals which are likely to carry cross-modal representations, it might be challenging to measure them without increasing their signal-to-noise ratio, for example, via a task instruction emphasizing the necessity to abstract away from modality-driven details. Our experimental task did not impose such necessity, since we were particularly interested in measuring automatic processing of the object category representations and their possible cross-modal generalization in the absence of task-driven effects. Instead, reading/listening or making high-level cognitive judgements (e.g., category attribution) would aid more conceptual stimuli assessment independent of modality-specific features. In line with this, other manipulations which allow to increase feedback signal-to-noise ratio or reduce the impact of feedforward signals potentially impeding the extraction of the top-down information (e.g., backward masking) could improve tracking of feedback contents during object recognition.

This work focused specifically on how feedforward and feedback signals activate visual representations within the visual system or across visual and auditory systems. However, at the moment the space of unknown regarding the organization of feedback signals and their interaction with feedforward inputs remains immense. There are at least three possible research directions that

could be pursued to unravel the mechanisms enabling feedforward/feedback information exchange in the brain:

(1) Does task-relevant feedback to V1 target the center of the visual field or visual periphery or both?

There are studies showing diverging results regarding the spatial distribution of high-level information across V1 cortical sheet: (1) enhanced feedback representations in the fovea compared to periphery; (2) enhanced feedback representation in the periphery compared to fovea; (3) homogeneous distribution of feedback information over the V1 space.

The first group of studies comprising work by Zhaoping and colleagues demonstrated that providing conflicting inputs in the left and right eye results in feature misbinding and visual illusions in the peripheral vision but not in the center of the visual field (for review see Zhaoping, 2019). These findings are interpreted as a result of task-relevant feedback from higher-order areas to V1 that majorly targets the foveal representation in order to facilitate object recognition and veto visual illusions. Thus, in this view task-related feedback signals prevail in the center of the visual field compared to visual periphery. This theoretical view (Zhaoping, 2019) is also supported by the results of the study (Williams et al., 2007), where participants assessed similarity of two objects shown in the visual periphery (more than 5 degrees outside of the fovea) but the object category information could be decoded specifically from within the foveal retinotopic cortex, presumably to enhance task performance. In direct contrast to the pattern of results showing enhanced feedback received by the fovea, the second group of studies demonstrated that the task-relevant categorical feedback about natural sounds (Vetter et al., 2014) and visual context feedback signals from hippocampus (Smith & Muckli, 2010) could be extracted mainly from peripheral early visual regions. Furthermore, the third group of studies shows that information about task-

relevant features (Serences & Boynton, 2007) could be extracted both from central and peripheral parts of the foveal retinotopic cortex. In line with this study a recent modeling work (Breedlove et al., 2020) showed that receptive fields of the feedback from higher processing levels (during mental imagery) are larger than those during perception possibly as a result of less sensitivity to stimulus variation in higher processing levels. Understanding how V1 receives and processes high-level feedback is essential to unravel how task-relevant features are incorporated to update the percept. In Study I and II of the present work, we extracted feedback signals from the foveal part of V1 (0-3 visual degrees) which is presumably at odds with the second line of research. However, future studies are needed to more systematically explore the distribution of task-relevant top-down features in the cortical space of V1 and identify factors which impact this distribution.

(2) Are feedforward and feedback effects driving and modulatory, respectively?

In this work, we followed a common view on feedback signals as exerting modulatory effects on sensory representations, which are, based on animal findings, assumed to be weaker than the feedforward input driving neuronal responses in V1 (Klink et al., 2017; Kerkoerle et al., 2017). One prominent reasoning to support this view is that V1 neurons, which further project to multiple downstream areas processing distinct input features, are not fully suppressed by the selective feedback from these downstream areas but keep supplying them with the feedforward input (Seidemann & Geisler, 2018). However, it has been shown in a series of studies (Covic & Sherman, 2011; De Pasquale & Sherman, 2011; Bastos et al., 2012) that modulatory and driving cells are present both among those projecting down- and upstream the visual hierarchy, hence making driving and modulatory effects at the level of neural populations orthogonal to the direction of the information stream. Driving feedback signals are considered essential, for example, for the predictive coding theory, since top-down predictions need to elicit responses in

their targets - cells which report prediction errors (Bastos et al., 2012). Future studies are needed to find out when feedback mediates modulatory and driving effects on sensory input (van Loon et al., 2016) and how it affects the contents of the feedback representations.

(3) How strongly feedforward/feedback information streams depend on the respective anatomical projections?

One of the premises of the current work was that abundance of feedback anatomical projections in the visual system necessitates massive recurrence in visual cognition. However, a recent study shows that cortico-cortical projections estimated over a hundred brain areas could predict the functional connectivity within frequency bands for feedforward - but not feedback - signal transmission (Vezoli et al., 2021). One of the suggested reasons for such a divergence between the feedback anatomical projections and the feedback functional connectivity is a possibility for the brain to flexibly establish functional connections between the areas depending on learning. Yet, the present empirical evidence (Jia et al., 2020) suggests the opposite: orientation discrimination learning in humans increases feedforward rather than feedback layer-to-layer connectivity possibly through gating of perceptual decisions via sensory plasticity. An exciting avenue for the future studies would be to investigate the effects of learning on employment of feedforward and feedback anatomical projections and establishing whether learning is the core factor improving the mapping between the feedback cortico-cortical projections and functional connectivity.

To summarize, this work focused on how feedforward and feedback signals are differentiated within the same sensory system and integrated across sensory systems. The results of our studies elucidate how perceived and internally-generated contents initiated by top-down signals are simultaneously represented in different cortical compartments of area V1. These results

reflect a general strategy for implementation of multiple cognitive functions within the same sensory system. For the more precise estimation of signal-by-depth separation in humans, we mapped out the choice space of paradigms and MR-protocols suitable for a typical cognitive neuroscience experiment thus further contributing to optimization of acquisition pipelines for depth-specific imaging at 7T fMRI. In this work, we did not find how sensory signals from one sensory system initiate visual representations and suggested experimental manipulations which could further test mechanisms for cross-modal signal transmission. With this, our results contribute to a large body of research interrogating how feedforward and feedback signfals give rise to complex visual cognition.

6 REFERENCES

Aitken F, Menelaou G, Warrington O, Koolschijn RS, Corbin N, et al. (2020) Prior expectations evoke stimulus-specific activity in the deep layers of the primary visual cortex. PLOS Biology 18(12): e3001023. https://doi.org/10.1371/journal.pbio.3001023

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & De Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. Current Biology, 23(15), 1427-1431.

Amedi, A., Raz, N., Pianka, P., Malach, R., & Zohary, E. (2003). Early 'visual' cortex activation correlates with superior verbal memory performance in the blind. Nature neuroscience, 6(7), 758-766.

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical image analysis, 12(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Bainbridge, W. A., Pounder, Z., Eardley, A. F., & Baker, C. I. (2021). Quantifying Aphantasia through drawing: Those without visual imagery show deficits in object but not spatial memory. Cortex, 135, 159-172.

Barone, P., Batardiere, A., Knoblauch, K., & Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *Journal of Neuroscience*, *20*(9), 3263-3281.

Bastos, A. M., Loonis, R., Kornblith, S., Lundqvist, M., & Miller, E. K. (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. *Proceedings of the National Academy of Sciences*, *115*(5), 1117-1122.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. Neuron, 76(4), 695-711.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., ... & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390-401.

Bause, J., Polimeni, J. R., Stelzer, J., In, M. H., Ehses, P., Kraemer-Fernandez, P., ... & Scheffler, K. (2020). Impact of prospective motion correction, distortion correction methods and large vein bias on the spatial accuracy of cortical laminar fMRI at 9.4 Tesla. Neuroimage, 208, 116434.

Beer, A. L., Plank, T., Meyer, G., & Greenlee, M. W. (2013). Combined diffusion-weighted and functional magnetic resonance imaging reveals a temporal-occipital network involved in auditory-visual object processing. *Frontiers in integrative neuroscience*, *7*, 5.

Benjamini, Y., & Yekutieli, D. (2005). False discovery rate–adjusted multiple confidence intervals for selected parameters. Journal of the American Statistical Association, 100(469), 71-81.

Benson, N. C., Butt, O. H., Brainard, D. H., & Aguirre, G. K. (2014). Correction of distortion in flattened representations of the cortical surface allows prediction of V1-V3 functional organization from anatomy. PLoS computational biology, 10(3), e1003538.

Bergmann, J., Morgan, A. T., & Muckli, L. Two distinct feedback codes in V1 for 'real' and 'imaginary' internal experiences. Preprint at: https://www.biorxiv.org/content/biorxiv/early/2019/06/13/664870.full.pdf., 664870 (2019).

Bijanzadeh, M., Nurminen, L., Merlin, S., Clark, A. M., & Angelucci, A. (2018). Distinct laminar processing of local and global context in primate primary visual cortex. Neuron, 100(1), 259-274.

Bosman, C. A., Schoffelen, J. M., Brunet, N., Oostenveld, R., Bastos, A. M., Womelsdorf, T., ... & Fries, P. (2012). Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron*, *75*(5), 875-888.

Boyacioglu R, Schulz J, Muller NC, Koopmans PJ, Barth M, Norris DG. Whole brain, high resolution multiband spin-echo EPI fMRI at 7 T: a comparison with gradient-echo EPI using a color-word Stroop task. Neuroimage 2014; 97: 142–150.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433-436.

Breedlove, J. L., St-Yves, G., Olman, C. A., & Naselaris, T. (2020). Generative feedback explains distinct brain activity codes for seen and mental images. *Current Biology*, *30*(12), 2211-2224.

Bullier J. 2001. Integrated model of visual processing. Brain Res. Brain Res. Rev. 36:96–107

Canales-Johnson, A., Lanfranco, R. C., Morales, J. P., Martínez-Pernía, D., Valdés, J., Ezquerro-Nassar, A., ... & Noreika, V. (2021). In your phase: neural phase synchronisation underlies visual imagery of faces. *Scientific reports*, *11*(1), 1-13.

Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience*, *22*(11), 2886-2902.

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: the first 1000 ms. *Journal of vision*, *13*(10), 1-1.

Chai, Y., Li, L., Huber, L., Poser, B. A., & Bandettini, P. A. (2020). Integrated VASO and perfusion contrast: A new tool for laminar functional MRI. NeuroImage, 207, 116358.

Christophel, T. B., & Haynes, J. D. (2014). Decoding complex flow-field patterns in visual working memory. *Neuroimage*, *91*, 43-51.

Christophel, T. B., Cichy, R. M., Hebart, M. N., & Haynes, J. D. (2015). Parietal and early visual cortices encode working memory content across mental transformations. Neuroimage, 106, 198-206.

Christophel, T. B.*, Iamshchinina, P.*, Yan, C., Allefeld, C., & Haynes, J. D. (2018). Cortical specialization for attended versus unattended working memory. *Nature neuroscience*, *21*(4), 494-496.

Cichy, R. M., Heinzle, J., & Haynes, J. D. (2012). Imagery and perception share cortical representations of content and location. *Cerebral cortex*, *22*(2), 372-380.

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature neuroscience*, *17*(3), 455-462.

Cooper, L. A. (1975). Mental rotation of random two-dimensional shapes. Cognitive psychology, 7(1), 20-43.

Cooper, L. A., & Podgorny, P. (1976). Mental transformations and visual comparison processes: Effects of complexity and similarity. Journal of Experimental Psychology: Human Perception and Performance, 2(4), 503–514.

Covic, E. N., & Sherman, S. M. (2011). Synaptic properties of connections between the primary and secondary auditory cortices in mice. *Cerebral Cortex*, *21*(11), 2425-2441.

De Pasquale, R., & Sherman, S. M. (2011). Synaptic properties of corticocortical connections between the primary and secondary visual cortical areas in the mouse. *Journal of Neuroscience*, *31*(46), 16494-16506.

Deniz, F., Nunez-Elizalde, A. O., Huth, A. G., & Gallant, J. L. (2019). The representation of semantic information across human cerebral cortex during listening versus reading is invariant to stimulus modality. *Journal of Neuroscience*, *39*(39), 7722-7736.

Dijkstra, N., Bosch, S. E., & van Gerven, M. A. (2019). Shared neural mechanisms of visual perception and imagery. Trends in cognitive sciences, 23(5), 423-434.

Dijkstra, N., Zeidman, P., Ondobaka, S., van Gerven, M. A., & Friston, K. (2017). Distinct top-down and bottom-up brain connectivity during visual perception and imagery. Scientific reports, 7(1), 1-9.

Duong, T. Q., Yacoub, E., Adriany, G., Hu, X., Uğurbil, K., & Kim, S. G. (2003). Microvascular BOLD contribution at 4 and 7 T in the human brain: gradient-echo and spin-echo fMRI with suppression of blood effects. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *49*(6), 1019-1027.

Eckert, M. A., Kamdar, N. V., Chang, C. E., Beckmann, C. F., Greicius, M. D., & Menon, V. (2008). A cross-modal system linking primary auditory and visual cortices: Evidence from intrinsic fMRI connectivity analysis. *Human brain mapping*, *29*(7), 848-857.

Ester, E. F., Sprague, T. C., & Serences, J. T. (2015). Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. *Neuron*, *87*(4), 893-905.

Fairhall, S. L., & Caramazza, A. (2013). Brain regions that represent amodal conceptual knowledge. *Journal of Neuroscience*, *33*(25), 10552-10558.

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., ... & Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PloS one*, *5*(12), e15710.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, *1*(1), 1-47.

Finn, E. S., Huber, L., Jangraw, D. C., Molfese, P. J., & Bandettini, P. A. (2019). Layer-dependent activity in human prefrontal cortex during working memory. *Nature neuroscience*, *22*(10), 1687-1695.

Fischler, I. (1985). Brain potentials during sentence verification: Automatic aspects of comprehension. *Biological Psychology, 21*(2), 83–105. doi:10.1016/0301-0511(85)90008-0

Fischler, I., Bloom, P. A., Childers, D. G., Arroyo, A. A., & Perry Jr, N. W. (1984). Brain potentials during sentence verification: Late negativity and long-term memory strength. *Neuropsychologia*, *22*(5), 559-568.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, *6*(2), 78-84.

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron*, *88*(1), 220-235.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, *291*(5508), 1560-1563.

Fujihara, N., Nageishi, Y., Koyama, S., & Nakajima, Y. (1998). Electrophysiological evidence for the typicality effect of human cognitive categorization. *International journal of psychophysiology*, *29*(1), 65-75.

Fulford, J., Milton, F., Salas, D., Smith, A., Simler, A., Winlove, C., & Zeman, A. (2018). The neural correlates of visual imagery vividness–An fMRI study and literature review. Cortex, 105, 26-40.

Gagnon, L., Sakadžić, S., Lesage, F., Pouliot, P., Dale, A. M., Devor, A., ... & Boas, D. A. (2016, April). Improving the calibrated fMRI estimation of CMRO2 with oxygen-sensitive Two-Photon Microscopy. In *Cancer Imaging and Therapy* (pp. JW3A-18). Optical Society of America.

Gagoski, B. A., Bilgic, B., Eichner, C., Bhat, H., Grant, P. E., Wald, L. L., & Setsompop, K. (2015). RARE/turbo spin echo imaging with simultaneous multislice Wave-CAIPI. *Magnetic resonance in medicine*, *73*(3), 929-938.

Giari, G., Leonardelli, E., Tao, Y., Machado, M., & Fairhall, S. L. (2020). Spatiotemporal properties of the neural representation of conceptual content for words and pictures–an MEG study. *Neuroimage*, *219*, 116913.

Goense, J. B., & Logothetis, N. K. (2006). Laminar specificity in monkey V1 using high-resolution SE-fMRI. *Magnetic resonance imaging*, *24*(4), 381-392.

Guggenmos, M., Sterzer, P., and Cichy, R. M. (2018). Multivariate pattern analysis for MEG: A comparison of dissimilarity measures. *NeuroImage*, *173*:434–447.

Haarsma, J., Kok, P., & Browning, M. (2020). The promise of layer-specific neuroimaging for testing predictive coding theories of psychosis. Schizophrenia Research.

Han, S., Eun, S., Cho, H., Uludağ, K., & Kim, S. G. (2021). Improvement of sensitivity and specificity for laminar BOLD fMRI with double spin-echo EPI in humans at 7 T. *NeuroImage*, *241*, 118435.

Harris, K. D., & Mrsic-Flogel, T. D. (2013). Cortical connectivity and sensory coding. Nature, 503(7474), 51–58. https://doi.org/10.1038/nature12654

Havlicek, M., & Uludağ, K. (2020). A dynamical model of the laminar BOLD response. *NeuroImage*, *204*, 116209.

Hebb, D. O. (1949). *The organisation of behaviour: a neuropsychological theory*. New York: Science Editions.

Heinzle, J., Koopmans, P. J., den Ouden, H. E., Raman, S., & Stephan, K. E. (2016). A hemodynamic model for layered BOLD signals. Neuroimage, 125, 556-570.

Hnazaee, M. F., & Van Hulle, M. M. (2017, May). Typicality effect on N400 ERP in categories despite differences in semantic processing. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 4379-4386). IEEE.

Horga, G., Schatz, K. C., Abi-Dargham, A., & Peterson, B. S. (2014). Deficits in predictive coding underlie hallucinations in schizophrenia. Journal of Neuroscience, 34(24), 8072-8082.

Huang, P., Correia, M. M., Rua, C., Rodgers, C. T., Henson, R. N., & Carlin, J. D. (2021). Correcting for Superficial Bias in 7T Gradient Echo fMRI. *Frontiers in Neuroscience*, 1170.

Huber, L., Ivanov, D., Handwerker, D. A., Marrett, S., Guidi, M., Uludağ, K., ... & Poser, B. A. (2018). Techniques for blood volume fMRI with VASO: From low-resolution mapping towards sub-millimeter layer-dependent applications. *Neuroimage*, *164*, 131-143.

Huber, L., Ivanov, D., Krieger, S. N., Streicher, M. N., Mildner, T., Poser, B. A., ... & Turner, R. (2014). Slab-selective, BOLD-corrected VASO at 7 Tesla provides measures of cerebral blood volume reactivity with high signal-to-noise ratio. *Magnetic resonance in medicine*, *72*(1), 137-148.

Huntenburg, J. M., Steele, C. J., & Bazin, P. L. (2018). Nighres: processing tools for high-resolution neuroimaging. GigaScience, 7(7), giy082.

Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453-458.

Huth, A. G., Nishimoto, S., Vu, A. T., & Gallant, J. L. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, *76*(6), 1210-1224.

Iamshchinina, P., Christophel, T. B., Gayet, S., & Rademaker, R. L. (2021a). Essential considerations for exploring visual working memory storage in the human brain. *Visual Cognition*, 1-12.

Iamshchinina, P., Kaiser, D., Yakupov, R., Haenelt, D., Sciarra, A., Mattern, H., ... & Cichy, R. M. (2021b). Perceived and mentally rotated contents are differentially represented in cortical depth of V1. *Commun Biol* **4,** 1069. https://doi.org/10.1038/s42003-021-02582-4

In, M. H., & Speck, O. (2012). Highly accelerated PSF-mapping for EPI distortion correction with improved fidelity. Magnetic Resonance Materials in Physics, Biology and Medicine, 25(3), 183-192.

Jia, K., Zamboni, E., Kemper, V., Rua, C., Goncalves, N. R., Ng, A. K. T., ... & Kourtzi, Z. (2020). Recurrent processing drives perceptual plasticity. *Current Biology*, *30*(21), 4177-4187.

Jung, Y., Larsen, B., & Walther, D. B. (2018). Modality-independent coding of scene categories in prefrontal cortex. *Journal of Neuroscience*, *38*(26), 5969-5981.

Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. Psychological review, 92(2), 137.

Kaiser, D., Azzalini, D. C., & Peelen, M. V. (2016). Shape-independent object category responses revealed by MEG and fMRI decoding. *Journal of neurophysiology*, *115*(4), 2246-2250.

Kar, K., & DiCarlo, J. J. (2021). Fast recurrent processing via ventrolateral prefrontal cortex is needed by the primate ventral stream for robust core visual object recognition. *Neuron*, *109*(1), 164-176.

Kashyap, S., Ivanov, D., Havlicek, M., Huber, L., Poser, B. A., & Uludağ, K. (2021). Sub-millimetre resolution laminar fMRI using Arterial Spin Labelling in humans at 7 T. *Plos one*, *16*(4), e0250504.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron, 22(4), 751-761.

Khooshabeh, P., Hegarty, M., & Shipley, T. F. (2013). Individual differences in mental rotation: Piecemeal versus holistic processing. Experimental psychology, 60(3), 164.

Klein, B. P., Fracasso, A., van Dijk, J. A., Paffen, C. L. E., te Pas, S. F., & Dumoulin, S. O. (2018). Cortical depth dependent population receptive field attraction by spatial attention in human V1. NeuroImage, 176, 301–312. https://doi.org/10.1016/j.neuroimage.2018.04.055

Klink, P. C., Dagnino, B., Gariel-Mathis, M. A., & Roelfsema, P. R. (2017). Distinct feedforward and feedback effects of microstimulation in visual cortex reveal neural mechanisms of texture segregation. *Neuron*, *95*(1), 209-220.

Kocagoncu, E., Clarke, A., Devereux, B. J., & Tyler, L. K. (2017). Decoding the cortical dynamics of sound-meaning mapping. *Journal of Neuroscience*, *37*(5), 1312-1319.

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*, *26*(3), 371-376.

Koopmans, P. J., Barth, M., & Norris, D. G. (2010). Layer-specific BOLD activation in human V1. *Human brain mapping*, *31*(9), 1297-1304.

Kumar, M., Federmeier, K. D., Fei-Fei, L., & Beck, D. M. (2017). Evidence for similar patterns of neural activity elicited by picture-and word-based representations of natural scenes. *Neuroimage*, *155*, 422-436.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, *4*(12), 463-470.

Lawrence, S. J., Norris, D. G., & De Lange, F. P. (2019). Dissociable laminar profiles of concurrent bottom-up and top-down modulation in the human visual cortex. Elife, 8, e44422.

Lawrence, S. J., van Mourik, T., Kok, P., Koopmans, P. J., Norris, D. G., & de Lange, F. P. (2018). Laminar organization of working memory signals in human visual cortex. Current Biology, 28(21), 3435-3440.

Lee, S. H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *Neuroimage*, *59*(4), 4064-4073.

Leonardelli, E., Fait, E., & Fairhall, S. L. (2019). Temporal dynamics of access to amodal representations of category-level conceptual information OPEN. *Scientific Reports, (December 2018), 1*–9.

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, *21*(6), 335-346.

Liu, C., Guo, F., Qian, C., Zhang, Z., Sun, K., Wang, D. J., He, S., & Zhang, P. (2020). Layer-dependent multiplicative effects of spatial attention on contrast responses in human early visual cortex. Progress in Neurobiology, 101897.

Lu, L., Wang, Q., Sheng, J., Liu, Z., Qin, L., Li, L., & Gao, J. H. (2019). Neural tracking of speech mental imagery during rhythmic inner counting. *Elife*, *8*, e48971.

Lund, J. S. (1988). Anatomical organization of macaque monkey striate visual cortex. Annual review of neuroscience, 11(1), 253-288.

Lüsebrink, F., Sciarra, A., Mattern, H., Yakupov, R., & Speck, O. (2017). T 1-weighted in vivo human whole brain MRI dataset with an ultrahigh isotropic resolution of 250 μm. *Scientific data*, *4*(1), 1-12.

Maclaren, J., Armstrong, B. S., Barrows, R. T., Danishad, K. A., Ernst, T., Foster, C. L., Gumus, K., Herbst, M., Kadashevich, I. Y., Kusik, T. P., Li, Q., Lovell-Smith, C., Prieto, T., Schulze, P., Speck, O., Stucht, D., & Zaitsev, M. (2012). Measurement and correction of microscopic head

motion during magnetic resonance imaging of the brain. PloS one, 7(11), e48088. https://doi.org/10.1371/journal.pone.0048088

Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods 164*, 177–190 (2007).

Markov, N. T., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., & Kennedy, H. (2013). Cortical high-density counterstream architectures. Science, 342(6158). https://doi.org/10.1126/science.1238406

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., … Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. Journal of Comparative Neurology, 522(1), 225–259. https://doi.org/10.1002/cne.23458

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., … Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. Journal of Comparative Neurology, 522(1), 225–259. https://doi.org/10.1002/cne.23458

Markuerkiaga, I., Barth, M., & Norris, D. G. (2016). A cortical vascular model for examining the specificity of the laminar BOLD signal. *Neuroimage*, *132*, 491-498.

Marquardt, I., Schneider, M., Gulban, O. F., Ivanov, D., & Uludağ, K. (2018). Cortical depth profiles of luminance contrast responses in human V1 and V2 using 7 T fMRI. *Human Brain Mapping*, *39*(7), 2812-2827.

Marques, J. P., & Norris, D. G. (2018). How to choose the right MR sequence for your research question at 7 T and above?. *NeuroImage*, *168*, 119-140.

Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004). Where bottom-up meets top-down: neuronal interactions during perception and imagery. Cerebral cortex, 14(11), 1256-1265.

Mejias, J. F., Murray, J. D., Kennedy, H., & Wang, X. J. (2016). Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. *Science advances*, *2*(11), e1601335.

Menon, R. S., Ogawa, S., & Uğurbil, K. (1995). High-temporal-resolution studies of the human primary visual cortex at 4 T: teasing out the oxygenation contribution in fMRI. *International Journal of Imaging Systems and Technology*, *6*(2-3), 209-215.

Michalareas, G., Vezoli, J., Van Pelt, S., Schoffelen, J. M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron*, *89*(2), 384-397.

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., 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.

Morgan, A. T., Petro, L. S., & Muckli, L. (2016). Cortical feedback to V1 and V2 contains unique information about high-level scene structure.

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., … Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. Current Biology, 25(20), 2690–2695. https://doi.org/10.1016/j.cub.2015.08.057

Mugler III, J. P. (2014). Optimized three-dimensional fast-spin-echo MRI. *Journal of magnetic resonance imaging*, *39*(4), 745-767.

Muller, K. R., Mika, S., Ratsch, G., Tsuda, K., & Scholkopf, B. (2001). An introduction to kernel-based learning algorithms. IEEE transactions on neural networks, 12(2), 181-201.

Musz, E., Loiotile, R. E., Chen, J., Cusack, R., & Bedny, M. (2021). Naturalistic stimuli reveal a critical period in visual cortex development: Evidence from adult-onset blindness. *bioRxiv*.

Nichols, T.E. & Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp. 15,* 1–25.

Norris, D. G., Koopmans, P. J., Boyacioğlu, R., & Barth, M. (2011). Power independent of number of slices (PINS) radiofrequency pulses for low-power simultaneous multislice excitation. *Magnetic resonance in medicine*, *66*(5), 1234-1240.

Núñez-Peña, M. I., & Honrubia-Serrano, M. L. (2005). N400 and category exemplar associative strength. *International Journal of Psychophysiology*, *56*(1), 45-54.

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of cognitive neuroscience*, *12*(6), 1013-1023.

Olman, C. A., Van de Moortele, P. F., Schumacher, J. F., Guy, J. R., Uğurbil, K., & Yacoub, E. (2010). Retinotopic mapping with spin echo BOLD at 7T. *Magnetic resonance imaging*, *28*(9), 1258-1269.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). Fieldtrip: open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience, 2011*:1.

Pearson, J. (2019). The human imagination: the cognitive neuroscience of visual mental imagery. Nature Reviews Neuroscience, 20(10), 624-634.

Persichetti, A. S., Avery, J. A., Huber, L., Merriam, E. P., & Martin, A. (2020). Layer-specific contributions to imagined and executed hand movements in human primary motor cortex. *Current Biology*, *30*(9), 1721-1725.

Polimeni, J. R., Fischl, B., Greve, D. N., & Wald, L. L. (2010). Laminar analysis of 7 T BOLD using an imposed spatial activation pattern in human V1. *Neuroimage*, *52*(4), 1334-1346.

Poort, J., Raudies, F., Wannig, A., Lamme, V. A., Neumann, H., & Roelfsema, P. R. (2012). The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. Neuron, 75(1), 143-156.

Popham, S. F., Huth, A. G., Bilenko, N. Y., Deniz, F., Gao, J. S., Nunez-Elizalde, A. O., & Gallant, J. L. (2021). Visual and linguistic semantic representations are aligned at the border of human visual cortex. *Nature Neuroscience*, *24*(11), 1628-1636.

Pounder, Z., Jacob, J., Jacobs, C., Loveday, C., Towell, T., & Silvanto, J. (2018). Mental rotation performance in aphantasia. Journal of Vision, 18(10), 1123-1123.

Pozzi, I., Bohte, S., & Roelfsema, P. (2020). Attention-Gated Brain Propagation: How the brain can implement reward-based error backpropagation. *Advances in Neural Information Processing Systems*, *33*.

Proklova, D., Kaiser, D., & Peelen, M. V. (2019). MEG sensor patterns reflect perceptual but not categorical similarity of animate and inanimate objects. *NeuroImage*, *193*, 167-177.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. Nature neuroscience, 22(8), 1336-1344. Hgfv

Roberts, J. E., & Bell, M. A. (2000). Sex differences on a mental rotation task: variations in electroencephalogram hemispheric activation between children and college students. Developmental neuropsychology, 17(2), 199-223.

Rockland, K. S., & Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. *International Journal of Psychophysiology*, *50*(1-2), 19-26.

Rockland, K. S., & Pandya, D. N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. Brain Research, 179(1), 3–20.

Rockland, K. S., & Pandya, D. N. (1979). Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. Brain Research, 179(1), 3–20.

Rockland, K. S., & Van Hoesen, G. W. (1994). Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cerebral cortex*, *4*(3), 300-313.

Roelfsema, P. R., & de Lange, F. P. (2016). Early visual cortex as a multiscale cognitive blackboard. Annual review of vision science, 2, 131-151.

Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. Nature, 395(6700), 376-381.

Rohenkohl, G., Bosman, C. A., & Fries, P. (2018). Gamma synchronization between V1 and V4 improves behavioral performance. *Neuron*, *100*(4), 953-963.

Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nature neuroscience*, *5*(7), 631-632.

Schmidt, T. T., & Blankenburg, F. (2019). The somatotopy of mental tactile imagery. *Frontiers in Human Neuroscience*, *13*, 10.

Seidemann, E., & Geisler, W. S. (2018). Linking V1 activity to behavior. *Annual review of vision science*, *4*, 287-310.

Self, M. W., van Kerkoerle, T., Goebel, R., & Roelfsema, P. R. (2019). Benchmarking laminar fMRI: neuronal spiking and synaptic activity during top-down and bottom-up processing in the different layers of cortex. Neuroimage, 197, 806-817.

Self, M. W., van Kerkoerle, T., Supèr, H., & Roelfsema, P. R. (2013). Distinct roles of the cortical layers of area V1 in figure-ground segregation. Current Biology, 23(21), 2121-2129.

Serences, J. T., & Boynton, G. M. (2007). Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, *55*(2), 301-312.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171(3972), 701-703.

Shinkareva, S. V, Malave, V. L., Mason, R. A., Mitchell, T. M., & Adam, M. (2011). Commonality of neural representations of words and pictures. *NeuroImage, 54*(3), 2418–2425.

Simanova, I., Hagoort, P., Oostenveld, R., and Van Gerven, M. A. (2012). Modality independent decoding of semantic information from the human brain. *Cerebral cortex*, *24*(2):426–434.

Simanova, I., Van Gerven, M., Oostenveld, R., and Hagoort, P. (2010). Identifying object categories from event-related eeg: toward decoding of conceptual representations. *PloS one, 5*(12):e14465.

Siu, C. R., Balsor, J. L., Merlin, S., Federer, F., & Angelucci, A. (2021). A direct interareal feedback-to-feedforward circuit in primate visual cortex. *BioRxiv*, 2020-07.

Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. Proceedings of the National Academy of Sciences, 107(46), 20099-20103.

Stauch, B. J., Peter, A., Schuler, H., & Fries, P. (2021). Stimulus-specific plasticity in human visual gamma-band activity and functional connectivity. *Elife*, *10*, e68240.

Stephan, K. E., Petzschner, F. H., Kasper, L., Bayer, J., Wellstein, K. V., Stefanics, G., ... & Heinzle, J. (2019). Laminar fMRI and computational theories of brain function. Neuroimage, 197, 699-706.

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., ... & Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological psychiatry*, *84*(9), 634-643.

Turner, R. (2002). How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes. *Neuroimage*, *16*(4), 1062-1067.

Turner, R. (2002). How much cortex can a vein drain? Downstream dilution of activation-related cerebral blood oxygenation changes. *Neuroimage*, *16*(4), 1062-1067.

Uludağ, K., & Blinder, P. (2018). Linking brain vascular physiology to hemodynamic response in ultra-high field MRI. Neuroimage, 168, 279-295.

Van de Moortele, P. F., Auerbach, E. J., Olman, C., Yacoub, E., Uğurbil, K., & Moeller, S. (2009). T1 weighted brain images at 7 Tesla unbiased for Proton Density, T2* contrast and RF coil receive B1 sensitivity with simultaneous vessel visualization. NeuroImage, 46(2), 432–446. https://doi.org/10.1016/j.neuroimage.2009.02.009

van der Zwaag, W., Schäfer, A., Marques, J. P., Turner, R., & Trampel, R. (2016). Recent applications of UHF-MRI in the study of human brain function and structure: a review. *NMR in Biomedicine*, *29*(9), 1274-1288.

Van Essen, D. C., & Felleman, D. J. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. Cerebral Cortex, 1, 1–47. Retrieved from http://www.cns.nyu.edu/~tony/vns/readings/felleman-vanessen-1991.pdf

van Kerkoerle, T., Self, M. W., & Roelfsema, P. R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. Nature Communications, 8(1), 13804. https://doi.org/10.1038/ncomms13804

Van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M. A., Poort, J., Van Der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. Proceedings of the National Academy of Sciences, 111(40), 14332-14341.

van Loon, A. M., Fahrenfort, J. J., van der Velde, B., Lirk, P. B., Vulink, N. C., Hollmann, M. W., ... & Lamme, V. A. (2016). NMDA receptor antagonist ketamine distorts object recognition by reducing feedback to early visual cortex. *Cerebral Cortex*, *26*(5), 1986-1996.

Vetter, P., Bola, Ł., Reich, L., Bennett, M., Muckli, L., & Amedi, A. (2020). Decoding natural sounds in early "visual" cortex of congenitally blind individuals. *Current Biology*, *30*(15), 3039-3044.

Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Current Biology*, *24*(11), 1256-1262.

Vezoli, J., Vinck, M., Bosman, C. A., Bastos, A. M., Lewis, C. M., Kennedy, H., & Fries, P. (2021). Brain rhythms define distinct interaction networks with differential dependence on anatomy. *Neuron*.

Waehnert, M. D., Dinse, J., Weiss, M., Streicher, M. N., Waehnert, P., Geyer, S., … Bazin, P. L. (2014). Anatomically motivated modeling of cortical laminae. NeuroImage, 93, 210–220. https://doi.org/10.1016/j.neuroimage.2013.03.078

Williams, M. A., Dang, S., & Kanwisher, N. G. (2007). Only some spatial patterns of fMRI response are read out in task performance. *Nature neuroscience*, *10*(6), 685-686.

Xie, S., Kaiser, D., & Cichy, R. M. (2020). Visual imagery and perception share neural representations in the alpha frequency band. *Current Biology*, *30*(13), 2621-2627.

Xu, Y. (2021). Towards a better understanding of information storage in visual working memory. *Visual Cognition*, *29*(7), 437-445.

Yacoub, E., Duong, T. Q., Van De Moortele, P. F., Lindquist, M., Adriany, G., Kim, S. G., ... & Hu, X. (2003). Spin-echo fMRI in humans using high spatial resolutions and high magnetic fields. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *49*(4), 655-664.

Yu, Y., Huber, L., Yang, J., Jangraw, D. C., Handwerker, D. A., Molfese, P. J., ... & Bandettini, P. A. (2019). Layer-specific activation of sensory input and predictive feedback in the human primary somatosensory cortex. *Science advances*, *5*(5), eaav9053.

Zaretskaya, N., Fischl, B., Reuter, M., Renvall, V., & Polimeni, J. R. (2018). Advantages of cortical surface reconstruction using submillimeter 7 T MEMPRAGE. *Neuroimage*, *165*, 11-26.

Zeman, A., Milton, F., Della Sala, S., Dewar, M., Frayling, T., Gaddum, J., ... & Winlove, C. (2020). Phantasia–the psychological significance of lifelong visual imagery vividness extremes. *Cortex*, *130*, 426-440.

Zhao, F., Wang, P., Hendrich, K., Ugurbil, K., & Kim, S. G. (2006). Cortical layer-dependent BOLD and CBV responses measured by spin-echo and gradient-echo fMRI: insights into hemodynamic regulation. *Neuroimage*, *30*(4), 1149-1160

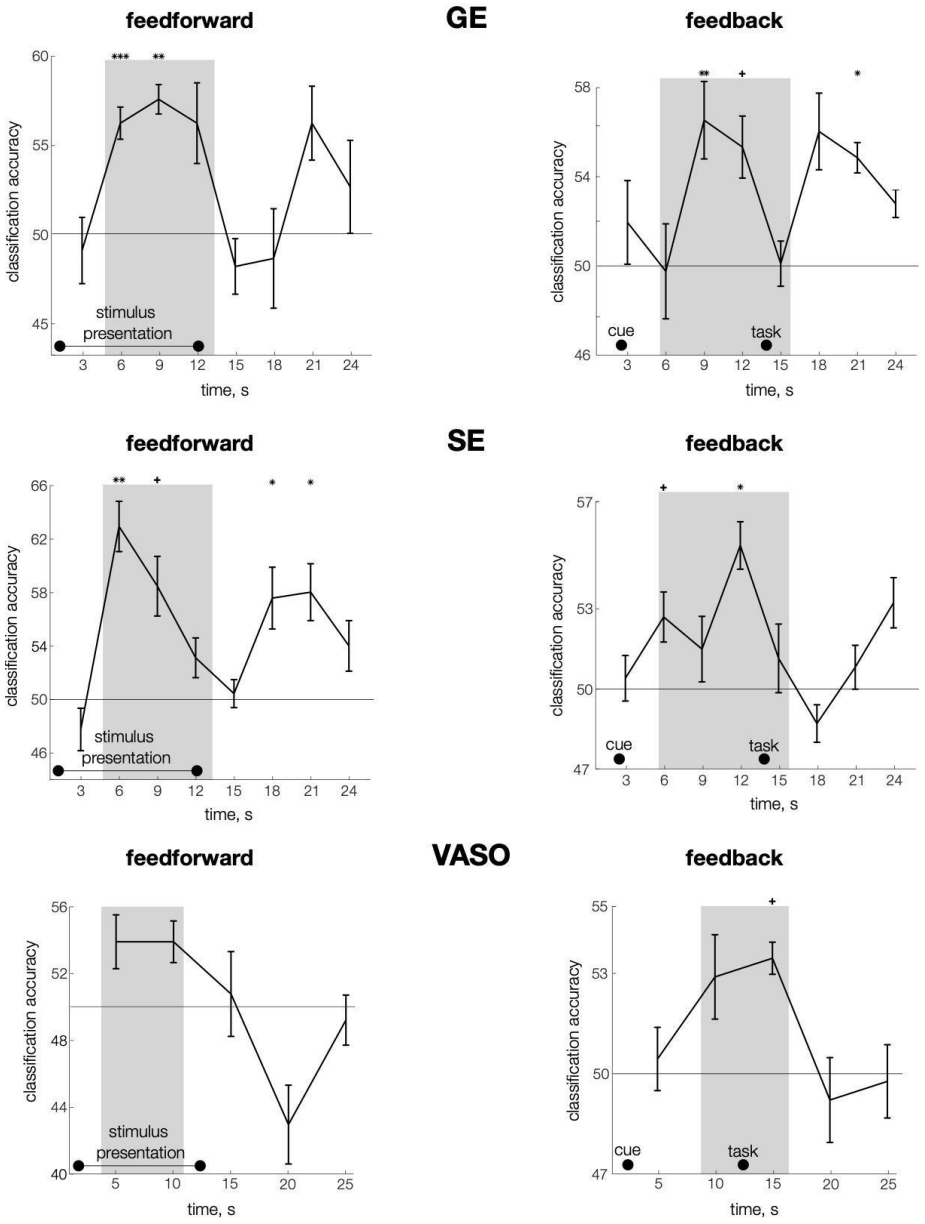
Zhaoping, L. (2019). A new framework for understanding vision from the perspective of the primary visual cortex. *Current Opinion in Neurobiology*, *58*, 1-10.

7 SUPPLEMENTARY MATERIALS

7.1 Supplementary Material – Study I

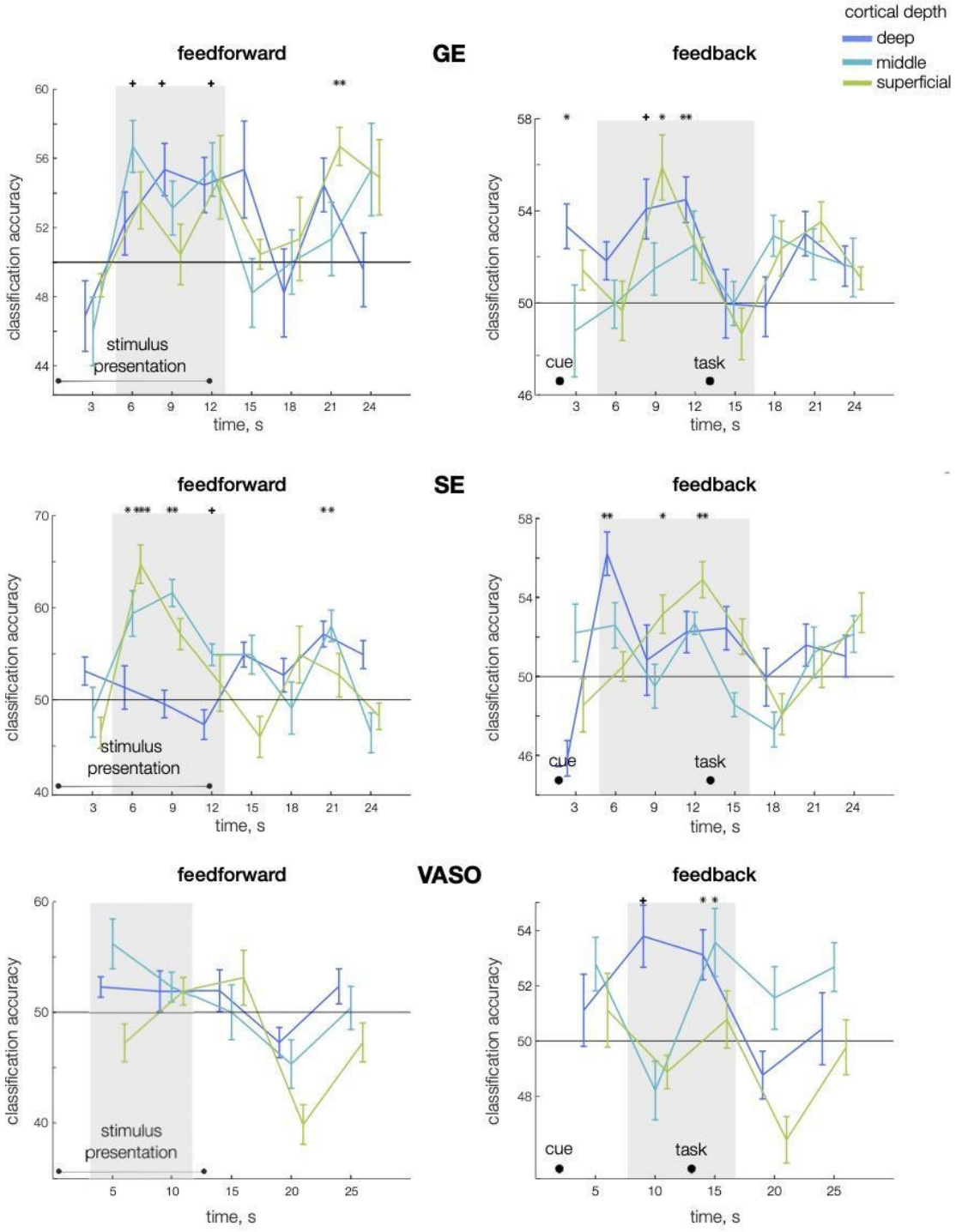
Supplementary materials for "Study I: Perceived and mentally rotated contents are differentially represented in cortical depth of V1" (pages 93-99) were removed for copyright reasons. The text can be found here: https://doi.org/10.1038/s42003-021-02582-4

7.2 Supplementary Material - Study II

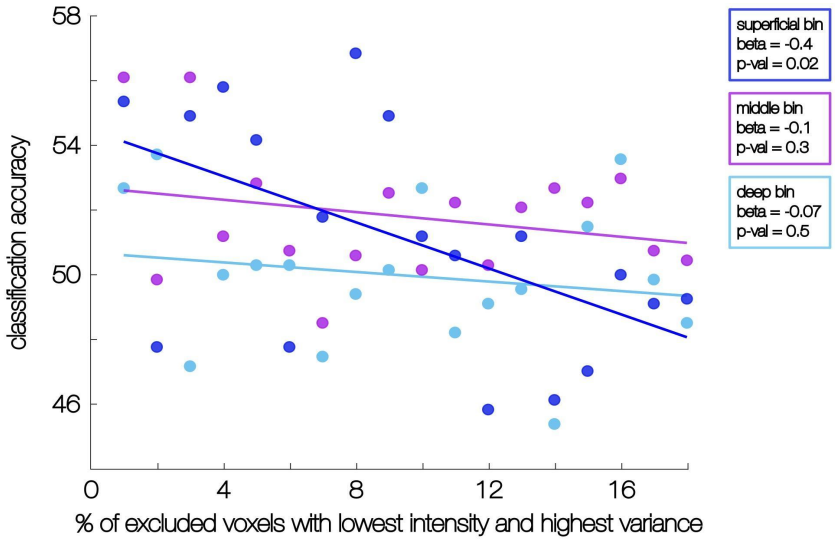


Supplementary Figure 1. Left column. Classification time series of feedforward (perceived) contents (i.e. grating orientations) extracted from the full grey matter ribbon (all voxels included in classification analysis) and represented for every MR sequence. During the stimulus presentation phase of the trial, a significant representation of perceived grating orientation emerged in the analysis of GE- and SE-acquired data ($t(26)=2.96$, $p=0.007$, $t(26)=3.1$, $p=0.004$ for GE and SE, respectively). VASO data yielded above-chance classification results for perceived orientation at the level of trend ($t(30)=1.9$, $p=0.07$). Thus, we established the presence of the feedforward signal in the data acquired with all

the sequences. Somewhat weaker representational strength obtained from VASO measurements could be due to coarser time resolution and consequently the inclusion of less time points in this analysis. **Right column.** Same as the left column but for the feedback (working memory) classification results. The analysis based on all the voxels in the V1 ROI showed an above-chance classification results for memorized orientation in all the sequences ($t(104)=1.9$, $p=0.06$, $t(102)=2.8$, $p=0.005$, $t(110)=2.6$, $p=0.01$ for GE, SE and VASO sequences, respectively). Thus, our results indicate that it is possible to decode working memory representations from the data acquired not only with GE 7T fMRI (Lawrence et al., 2018) but also SE-EPI and VASO scanning methods.



Supplementary Figure 2. Left column. Time series of classification accuracy for feedforward contents (perceived orientation) extracted from the three approximated cortical depth bins and represented for every MR sequence. **Right column.** Same as the left column but for the feedback contents (memorized orientation).



Supplementary Figure 3. Exclusion of the voxels potentially containing input from vein vessels in the data acquired during perception task with SE-EPI sequence. Voxels with the lowest intensity and highest variance (criteria taken from Menson et al., 1993) were excluded before performing the classification analysis at different percentages (from 2% to 19% of voxels, x-axis). We assumed that if voxels at the superficial cortical level contain macrovasculature, the exclusion of these voxels should decrease the classification accuracy in the superficial bin but not in the middle or deep bins. Fitting linear slopes in the classifier estimates obtained from each cortical bin revealed a significant decrease in decoding from the superficial compartment (blue, slope significantly different from 0 slope, p=0.02), as more voxels were excluded. Such a tendency was not observed in the middle (pink) and deep (cyan) cortical bins (p=0.3 and p=0.5, respectively).

8 APPENDIX

8.1 List of Publications

1. Iamshchinina, P., Kaiser, D., Yakupov, R., Haenelt, D., Sciarra, A., Mattern, H., ... & Cichy, R. M. (2021). Perceived and mentally rotated contents are differentially represented in cortical depth of V1. *Commun Biol* **4,** 1069. https://doi.org/10.1038/s42003-021-02582-4
2. Iamshchinina, P., Haenelt, D., Trampel, R., Weisskopf, N., Kaiser, D., & Cichy, R. M. (2021). Benchmarking GE-EPI, SE-EPI and SS-SI-VASO sequences for depth-dependent separation of feedforward and feedback signals in high-field MRI. *bioRxiv*.
3. Iamshchinina, P., Karapetian, A., Kaiser, D., & Cichy, R. M. (2021). Resolving the time course of visual and auditory object categorization. *bioRxiv*.

8.2 Summary of the Main Results

Vast majority of visual cognitive functions from low to high level rely not only on feedforward signals carrying sensory input to downstream brain areas but also on internally-generated feedback signals traversing the brain in the opposite direction. The feedback signals underlie our ability to conjure up internal representations regardless of sensory input – when imagining an object or directly perceiving it. Despite ubiquitous implications of feedback signals in visual cognition, little is known about their functional organization in the brain.

Multiple studies have shown that within the visual system the same brain region can concurrently represent feedforward and feedback contents. Given this spatial overlap, **(1) how does the visual brain separate feedforward and feedback signals thus avoiding a mixture of the perceived and the imagined?** Confusing the two information streams could lead to potentially detrimental consequences. Another body of research demonstrated that feedback connections between two different sensory systems participate in a rapid and effortless signal transmission across them. **(2) How do nonvisual signals elicit visual representations?**

In this work, we aimed to scrutinize the functional organization of directed signal transmission in the visual brain by interrogating these two critical questions. In Studies I and II, we explored the functional segregation of feedforward and feedback signals in grey matter depth of early visual area V1 using 7T fMRI. In Study III we investigated the mechanism of cross-modal generalization using EEG.

In **Study I**, we hypothesized that functional segregation of external and internally-generated visual contents follows the organization of feedforward and feedback anatomical projections revealed in primate tracing anatomy studies: feedforward projections were found to

terminate in the middle cortical layer of primate area V1, whereas feedback connections project to the superficial and deep layers. We used high-resolution layer-specific fMRI and multivariate pattern analysis to test this hypothesis in a mental rotation task. We found that rotated contents were predominant at outer cortical depth compartments (i.e. superficial and deep). At the same time perceived contents were more strongly represented at the middle cortical compartment. These results correspond to the previous neuroanatomical findings and identify how through cortical depth compartmentalization V1 functionally segregates rather than confuses external from internally-generated visual contents.

For the more precise estimation of signal-by-depth separation revealed in Study I, next we benchmarked three MR-sequences at 7T - gradient-echo, spin-echo, and vascular space occupancy - in their ability to differentiate feedforward and feedback signals in V1. The experiment in **Study II** consisted of two complementary tasks: a perception task that predominantly evokes feedforward signals and a working memory task that relies on feedback signals. We used multivariate pattern analysis to read out the perceived (feedforward) and memorized (feedback) grating orientation from neural signals across cortical depth. Analyses across all the MR-sequences revealed perception signals predominantly in the middle cortical compartment of area V1 and working memory signals in the deep compartment. Despite an overall consistency across sequences, spin-echo was the only sequence where both feedforward and feedback information were differently pronounced across cortical depth in a statistically robust way. We therefore suggest that in the context of a typical cognitive neuroscience experiment manipulating feedforward and feedback signals at 7T fMRI, spin-echo method may provide a favorable trade-off between spatial specificity and signal sensitivity.

In **Study III** we focused on the second critical question - how are visual representations activated by signals belonging to another sensory modality? Here we built our hypothesis following the studies in the field of object recognition, which demonstrate that abstract category-level representations emerge in the brain after a brief stimuli presentation in the absence of any explicit categorization task. Based on these findings we assumed that two sensory systems can reach a modality-independent representational state providing a universal feature space which can be read out by both sensory systems. We used EEG and a paradigm in which participants were presented with images and spoken words while they were conducting an unrelated task. We aimed to explore whether categorical object representations in both modalities reflect a convergence towards modality-independent representations. We obtained robust representations of objects and object categories in visual and auditory modalities; however, we did not find a conceptual representation shared across modalities at the level of patterns extracted from EEG scalp electrodes in our study.

Overall, our results show that feedforward and feedback signals are spatially segregated in the grey matter depth, possibly reflecting a general strategy for implementation of multiple cognitive functions within the same brain region. This differentiation can be revealed with diverse MR-sequences at 7T fMRI, where spin-echo sequence could be particularly suitable for establishing cortical depth-specific effects in humans. We did not find modality-independent representations which, according to our hypothesis, may subserve the activation of visual representations by the signals from another sensory system. This pattern of results indicates that identifying the mechanisms bridging different sensory systems is more challenging than exploring within-modality signal circuitry and this challenge requires further studies. With this, our results

contribute to a large body of research interrogating how feedforward and feedback signals give rise to complex visual cognition.

8.3 Kurzfassungen der Ergebnisse

Ein Großteil der visuell-kognitiven Funktionen, auf niedrigen und höheren Ebenen, beruht nicht nur auf Feedforward-Signalen, die sensorische Eingaben zu nachgeschalteten Gehirnbereichen übertragen, sondern auch auf intern erzeugten Feedback-Signalen, die das Gehirn in die entgegengesetzte Richtung durchqueren. Die Feedback-Signale liegen unserer Fähigkeit zugrunde, innere Repräsentationen aufzurufen, unabhängig von sensorischen Eingaben – bei der Vorstellung eines Objekts oder der direkten Wahrnehmung. Trotz allgegenwärtiger Implikationen von Feedback-Signalen in der visuellen Wahrnehmung ist wenig über ihre funktionelle Organisation im Gehirn bekannt.

Mehrere Studien haben gezeigt, dass innerhalb des visuellen Systems dieselbe Hirnregion gleichzeitig Feedforward- und Feedback-Inhalte darstellen kann. Angesichts dieser räumlichen Überlappung (1) wie trennt das visuelle Gehirn Feedforward- und Feedback-Signale und vermeidet so eine Vermischung des Wahrgenommenen und des Imaginierten? Eine Verwechslung der beiden Informationsströme könnte potenziell nachteilige Folgen haben. Eine andere Forschungsgruppe zeigte, dass Feedback-Verbindungen zwischen zwei verschiedenen sensorischen Systemen an einer schnellen und mühelosen Signalübertragung über sie beteiligt sind. (2) Wie rufen nicht-visuelle Signale visuelle Repräsentationen hervor?

In dieser Arbeit untersuchten wir die funktionelle Organisation der gerichteten Signalübertragung im visuellen Gehirn, indem wir diese beiden kritischen Fragen näher betrachteten. In den Studien I und II untersuchten wir die funktionelle Trennung von Feedforward- und Feedback-Signalen in der Tiefe der grauen Substanz des frühen visuellen Bereichs V1 mithilfe

von 7T-fMRT. In Studie III untersuchten wir den Mechanismus der modalen Verallgemeinerung mittels EEG.

In Studie I stellten wir die Hypothese auf, dass die funktionelle Trennung von extern und intern generierten visuellen Inhalten der Organisation von anatomischen Feedforward- und Feedback-Projektionen folgt, die in Anatomiestudien von Primaten gezeigt wurden: Es wurde festgestellt, dass die Feedforward-Projektionen in der mittleren kortikalen Schicht des Primatenbereichs V1 enden, während Feedback-Verbindungen in die oberflächlichen und tiefen Schichten projizieren. Wir haben hochauflösende schichtspezifische fMRT und eine multivariate Musteranalyse verwendet, um diese Hypothese in einer mentalen Rotationsaufgabe zu testen. Wir fanden heraus, dass rotierte Inhalte in den äußeren kortikalen Tiefenkompartimenten (d. h. oberflächlich und tief) vorherrschen. Gleichzeitig waren die wahrgenommenen Inhalte stärker im mittleren kortikalen Kompartiment vertreten. Diese Ergebnisse entsprechen den bisherigen neuroanatomischen Befunden und zeigen auf, wie V1 durch die kortikale Tiefenkompartimentierung funktionell getrennte visuelle Inhalte von intern erzeugten Inhalten trennt, anstatt sie zu verwechseln.

Für die genauere Schätzung der Signal-durch-Tiefe-Trennung, die in Studie I gezeigt wurde, haben wir als nächstes drei MR-Sequenzen bei 7T – Gradienten echo, Spin-Echo und Gefäßraumbelegung – hinsichtlich ihrer Fähigkeit, Feedforward- und Feedback-Signale in V1 zu unterscheiden. Das Experiment in Studie II bestand aus zwei sich ergänzenden Aufgaben: einer Wahrnehmungsaufgabe, die überwiegend Feedforward-Signale hervorruft, und einer Arbeitsgedächtnisaufgabe, die auf Feedback-Signalen beruht. Wir verwendeten eine multivariate Musteranalyse, um die wahrgenommene (Feedforward) und gespeicherte (Feedback)

Gitterorientierung aus neuronalen Signalen über die kortikale Tiefe auszulesen. Analysen über alle MR-Sequenzen hinweg ergaben Wahrnehmungssignale überwiegend im mittleren kortikalen Kompartiment des Bereichs V1 und Arbeitsgedächtnissignale im tiefen Kompartiment. Trotz einer Gesamtkonsistenz über die Sequenzen hinweg war Spin-Echo die einzige Sequenz, bei der sowohl Feedforward- als auch Feedback-Informationen statistisch robust über die kortikale Tiefe unterschiedlich ausgeprägt waren. Wir schlagen daher vor, dass im Kontext eines typischen kognitiven neurowissenschaftlichen Experiments, bei dem Feedforward- und Feedback-Signale bei 7T fMRT manipuliert werden, die Spin-Echo-Methode einen vorteilhaften Kompromiss zwischen räumlicher Spezifität und Signalempfindlichkeit bieten kann.

In Studie III haben wir uns auf die zweite kritische Frage konzentriert – wie werden visuelle Repräsentationen durch Signale aktiviert, die zu einer anderen Sinnesmodalität gehören? Hier haben wir unsere Hypothese auf Studien des Gebiets der Objekterkennung aufgebaut, die zeigen, dass abstrakte Repräsentationen auf Kategorieebene im Gehirn nach einer kurzen Reizpräsentation ohne explizite Kategorisierungsaufgabe entstehen. Basierend auf diesen Erkenntnissen nahmen wir an, dass zwei sensorische Systeme einen modalitätsunabhängigen Repräsentationszustand erreichen können, der einen universellen Merkmalsraum bereitstellt, der von beiden sensorischen Systemen ausgelesen werden kann. Wir verwendeten EEG und ein Paradigma, bei dem den Teilnehmenden Bilder und gesprochene Worte präsentiert wurden, während sie eine dazu unabhängige Aufgabe durchführten. Wir wollten untersuchen, ob kategoriale Objektrepräsentationen in beiden Modalitäten eine Konvergenz zu modalitätsunabhängigen Repräsentationen widerspiegeln. Wir erhielten robuste Darstellungen von Objekten und Objektkategorien in visuellen und auditiven Modalitäten; konnten in unserer Studie jedoch keine

konzeptionelle Darstellung finden, die auf der Ebene der Muster, die von EEG-Kopfhautelektroden extrahiert wurden, über alle Modalitäten hinweg geteilt wurde.

Insgesamt zeigen unsere Ergebnisse, dass Feedforward- und Feedback-Signale in der Tiefe der grauen Substanz räumlich getrennt sind, was möglicherweise eine allgemeine Strategie zur Implementierung mehrerer kognitiver Funktionen innerhalb derselben Hirnregion widerspiegelt. Diese Differenzierung lässt sich mit diversen MR-Sequenzen bei 7T fMRT aufzeigen, wobei Spin-Echo-Sequenzen besonders geeignet sein könnten, um kortikale Tiefen-spezifische Effekte beim Menschen nachzuweisen. Wir fanden keine modalitätsunabhängigen Repräsentationen, die unserer Hypothese zufolge der Aktivierung visueller Repräsentationen durch die Signale eines anderen Sinnessystems dienen könnten. Dieses Ergebnismuster weist darauf hin, dass die Identifizierung der Mechanismen, die verschiedene sensorische Systeme überbrücken, schwieriger ist als die Untersuchung von Signalschaltungen innerhalb einer Modalität, und diese Herausforderung erfordert weitere Studien. Damit tragen unsere Ergebnisse zu einer umfangreichen Forschungsarbeit bei, die untersucht, wie Feedforward- und Feedback-Signale zu komplexer visueller Wahrnehmung führen.

8.4 Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt,

- dass ich die vorliegende Arbeit eigenständig und ohne unerlaubte Hilfe verfasst habe,
- dass Ideen und Gedanken aus Arbeiten anderer entsprechend gekennzeichnet wurden,
- dass ich mich nicht bereits anderwärtig um einen Doktorgrad beworben habe und keinen Doktorgrad in dem Promotionsfach Psychologie besitze, sowie
- dass ich die zugrundeliegende Promotionsordnung vom 08.08.2016 anerkenne

Berlin, 22.11.2021

Polina Iamshchinina

8.5 Eigenanteil

Erklärung gemäß § 7 Abs. 3 Satz 4 der Promotionsordnung über den Eigenanteil an den veröffentlichten oder zur Veröffentlichung vorgesehenen eingereichten wissenschaftlichen Schriften im Rahmen meiner publikationsbasierten Arbeit

I. Name, Vorname: Iamshchinina, Polina

Institut: Fachbereich Erziehungswissenschaft und Psychologie

Promotionsfach: Psychologie

Titel: Master of Science (MSc)

II. Nummerierte Aufstellung der eingereichten Schriften (Titel, Autoren, wo und wann veröffentlicht bzw. eingereicht):

1. Iamshchinina, P., Kaiser, D., Yakupov, R., Haenelt, D., Sciarra, A., Mattern, H., ... & Cichy, R. M. (2021). Perceived and mentally rotated contents are differentially represented in cortical depth of V1. *Commun Biol* **4,** 1069. https://doi.org/10.1038/s42003-021-02582-4

2. Iamshchinina, P., Haenelt, D., Trampel, R., Weisskopf, N., Kaiser, D., & Cichy, R. M. (2021). Benchmarking GE-EPI, SE-EPI and SS-SI-VASO sequences for depth-dependent separation of feedforward and feedback signals in high-field MRI. Manuscript submitted for publication in *NeuroImage*.

3. Iamshchinina, P., Karapetian, A., Kaiser, D., & Cichy, R. M. (2021). Resolving the time course of visual and auditory object categorization. Manuscript submitted for publication in *Journal of Neurophysiology (JNP)*.

III. Darlegung des eigenen Anteils der Schriften:

Die Bewertung des Eigenanteils entspricht der Skala: “vollständig – überwiegend – mehrheitlich – in Teilen”.

Zu II.1.: Konzeption (überwiegend), Versuchsdesign (überwiegend), Programmierung (vollständig), Datenerhebung (mehrheitlich), Datenanalyse (überwiegend), Datenauswertung (überwiegend), Erstellen des Manuskriptes (mehrheitlich).

Zu II.2.: Konzeption (überwiegend), Versuchsdesign (vollständig), Programmierung (vollständig), Datenerhebung (mehrheitlich), Datenanalyse (vollständig), Datenauswertung (vollständig), Erstellen des Manuskriptes (überwiegend).

Zu II.3.: Konzeption (überwiegend), Versuchsdesign (vollständig), Programmierung (vollständig), Datenerhebung (mehrheitlich), Datenanalyse (vollständig), Datenauswertung (in Teilen), Erstellen des Manuskriptes (überwiegend).

IV. Die Namen und Anschriften nebst E-Mail oder Fax der jeweiligen Mitautorinnen oder Mitautoren:

zu II.1.:

Daniel Kaiser, Mathematical Institute, Department of Mathematics and Computer Science, Physics, Geography, Justus-Liebig-Universität Gießen, Arndtstraße 2, 35392 Gießen; Center for Mind, Brain and Behavior (CMBB), Philipps-Universität Marburg and Justus-Liebig-Universität Gießen, Hans-Meerwein-Straße 6, 35032 Marburg, Germany; danielkaiser.net@gmail.com

Renat Yakupov, German Center for Neurodegenerative Diseases (DZNE), Leipziger Straße 44, Haus 64, Magdeburg 39120, Germany; Renat.Yakupov@dzne.de

Daniel Haenelt, Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, Leipzig 04103, Germany; haenelt@cbs.mpg.de

Alessandro Sciarra, Department of Biomedical Magnetic Resonance, Institute for Physics, Otto-von-Guericke-University, Universitätsplatz 2, Magdeburg 39106, Germany; Department of Neurology, Otto-von-Guericke University, Leipziger Str. 44, Magdeburg 39120, Germany; alessandro.sciarra@med.ovgu.de

Hendrik Mattern, Department of Biomedical Magnetic Resonance, Institute for Physics, Otto-von-Guericke-University, Universitätsplatz 2, Magdeburg 39106, Germany; hendrik.mattern@ovgu.de

Falk Luesebrink, Department of Biomedical Magnetic Resonance, Institute for Physics, Otto-von-Guericke-University, Universitätsplatz 2, Magdeburg 39106, Germany; Department of Neurology, Otto-von-Guericke University, Leipziger Str. 44, Magdeburg 39120, Germany; falk.luesebrink@med.ovgu.de

Emrah Duezel, German Center for Neurodegenerative Diseases (DZNE), Leipziger Straße 44, Haus 64, Magdeburg 39120, Germany; Center for Behavioral Brain Sciences, Universitätsplatz 2, G24-205, Magdeburg D-39106, Germany; emrah.duezel@dzne.de

Oliver Speck, German Center for Neurodegenerative Diseases (DZNE), Leipziger Straße 44, Haus 64, Magdeburg 39120, Germany; Department of Biomedical Magnetic Resonance, Institute for Physics, Otto-von-Guericke-University, Universitätsplatz 2, Magdeburg 39106, Germany; Center for Behavioral Brain Sciences, Universitätsplatz 2, G24-205, Magdeburg D-39106, Germany; Leibniz Institute for Neurobiology, Brenneckestraße 6, Magdeburg 39118, Germany; oliver.speck@ovgu.de

Nikolaus Weiskopf, Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain

Sciences, Stephanstraße 1a, Leipzig 04103, Germany; Felix Bloch Institute for Solid State Physics, Faculty of Physics and Earth Sciences, Leipzig University, Linnéstraße 5, 04103 Leipzig, Germany; weiskopf@cbs.mpg.de

Radoslaw Martin Cichy, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, Berlin 14195, Germany; Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Unter den Linden 6, Berlin 10099, Germany; rmcichy@zedat.fu-berlin.de

zu II.2.:

Daniel Kaiser, s.o.

Daniel Haenelt, s.o.

Robert Trampel, Department of Neurophysics, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, Leipzig 04103, Germany; trampel@cbs.mpg.de

Nikolaus Weiskopf, s.o.

Radoslaw Martin Cichy, s.o.

zu II.3.:

Agnessa Karapetian, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, Berlin 14195, Germany; agnessakarapetian@gmail.com

Daniel Kaiser, s.o.

Radoslaw Martin Cichy, s.o.

Ich bestätige die von Polina Iamshchinina unter III. angegebene Erklärung:

Name: **Daniel Kaiser** Unterschrift:_____
Name: **Renat Yakupov** Unterschrift:_____
Name: **Daniel Haenelt** Unterschrift:_____
Name: **Alessandro Sciarra** Unterschrift: _____
Name: **Hendrik Mattern** Unterschrift:_____
Name: **Falk Luesebrink** Unterschrift:_____
Name: **Emrah Duezel** Unterschrift:_____
Name: **Oliver Speck** Unterschrift:_____
Name: **Nikolaus Weiskopf** Unterschrift: _____
Name: **Radoslaw Martin Cichy** Unterschrift:_____
Name: **Agnessa Karapetian** Unterschrift:_____
Name: **Robert Trampel** Unterschrift:_____