Research MSc. in Cognitive Neuroscience (Perception, Action and Control Specialization)

Donders Graduate School of Cognitive Neuroscience

MSc. Thesis

Laminar fMRI at 3T: A replication attempt of top-down and bottom-up laminar activity in the primary visual cortex

J. Karolis Degutis, Benedikt V. Ehinger, Floris P. de Lange

Student name: J. Karolis Degutis

Student number: s1023026

Date: 03/08/2020

Primary Supervisor: Benedikt Ehinger

First reader: Floris P. de Lange

Second reader: David Norris

Table of Contents

Abstract

The emerging field of laminar functional MRI (fMRI) has the capability of non-invasively measuring depth-dependent activity within regions of the neocortex. Most previous studies using laminar fMRI have been done at ultra-high field (7T and above) as it allowed for higher spatial resolution; yet, recent innovations in MR sequences have enabled submillimeter functional resolution at 3T. This study aimed to replicate two previous 7T laminar fMRI findings while using a 3T MRI scanner and thus tested whether laminar fMRI could be used for cognitively-relevant research questions when scanning at a lower field strength. A data quality comparison between the current findings and a previous 7T dataset (Lawrence et al. 2018) found lower tSNR and lower functional t-contrast activation at 3T. Additionally, the study failed to replicate the two previous results: the bottom-up stimulus-contrast yielded non-specific agranular compared to granular layer activation, while the top-down visual working memory main effect could not be replicated due to lower functional t-contrasts. This study provides evidence against the feasibility of laminar fMRI at 3T.

Introduction

Areas within the brain communicate using top-down (feedback) and bottom-up (feedforward) signals. A simplified distinction between these two types of signal pathways is based on their anatomical area of origin. Feedforward signals originate in sensory organs and propagate from lower-sensory areas into higher-sensory areas, while feedback signals propagate the other way - from higher sensory areas to lower ones. Both feedback and feedforward signals have been shown to carry differential information across the brain (Barbas, 2015; Van Essen et al., 2001). For example, previous studies have distinguished between top-down signals for imagined stimuli and bottom-up signals for physically perceived stimuli (Albers, Kok, Toni, Dijkerman, & De Lange, 2013). A given brain region can be a target of both bottom-up and top-down signals, and these signal inputs arrive in anatomically distinct parts within the region. More precisely, the bottom-up and top-down inputs can be segregated based on which gray matter layer the signal arrives in.

Layers within the gray matter of a region make up a fundamental building block of the brain. Layers receive inputs and propagate information either within the region itself or to other connected areas (Barbas, 2015). As mentioned, when examining the inputs of a region more closely, a distinct association can be formed between feedforward or feedback inputs and which layer they arrive in. As an example, the human primary visual cortex (V1) has six histological layers (Balaram, Young, & Kaas, 2014). The anatomical feedback and feedforward circuits show that bottom-up information from the lateral geniculate nucleus, a signal which predominantly originated in the eye, primarily arrives in layer 4b, while modulatory top-down signals from the extrastriate visual cortices, parietal, and even frontal regions enter layer 2 and 5-6 (Callaway, 1998). Thus, there is a spatial disassociation in the input of V1: superficial and deep layers (also known as agranular layers) receive feedback signals, while the middle layers (granular layers) receive feedforward signals.

Up until recently, measuring feedback and feedforward signals within humans non-invasively was difficult. Thus, previous studies investigating how layers relate to top-down and bottom-up functional signals relied on invasive techniques within primates (Callaway, 1998). However, an increase in the spatial resolution of functional magnetic resonance imaging (fMRI) has enabled the examination of this fine-grained information flow within

healthy human participants. The nascent field of laminar fMRI uses ultrahigh spatial resolution to measure separable signals from different depths of the neocortex. Researchers have already looked at cognitive task laminar profiles within regions such as the early visual cortices (Kok, Bains, van Mourik, Norris, & de Lange, 2016; Lawrence, van Mourik, Kok, Koopmans & de Lange 2018; Lawrence, Norris, & De Lange, 2019), the motor cortex (Huber et al., 2017), language regions (Sharoh, van Mourik, Bains, Segaert, K., Hagoort, & Norris, 2019) and even the prefronal cortex (Finn, Huber, Jangraw, & Bandettini, 2018).

Most laminar fMRI research, including the aforementioned studies, used ultra-high field MRI of 7T and above, yet research looking into the possibility of extending laminar fMRI to lower static magnetic fields has gained attention. Previous studies have resorted to ultra-high field scanners, as it enabled an increase of spatial resolution; however, recent developments within scanning sequences (and new MR gradient coil capabilities) have allowed for functional submillimeter isotropic resolution at 3T (Koopmans, Barth, & Norris, 2010; Markuerkiaga, Marques, Bains, & Norris, 2020). A higher number of labs see the potential of investigating functional networks between brain regions while using a readily available MRI, as seen from a recent study which combined EEG and 3T fMRI to investigate how laminar activation profiles of V1 correspond to alpha and gamma band oscillations (Scheeringa, Koopmans, van Mourik, Jensen, & Norris, 2016). Despite a few studies using 3T for laminar fMRI, there still has not been one which investigates whether cognitively relevant laminar activity can be replicated at 3T. Such a replication would provide evidence for the feasibility of using a 3T scanner for novel laminar fMRI research.

The current study aimed to replicate two previous distinct laminar fMRI effects at 3T: a top-down laminar effect from Lawrence et al. 2018 using a visual working memory task and a bottom-up laminar effect from Lawrence et al. 2019 using a stimulus-contrast task. Additionally, we compared the data quality between 3T and 7T datasets.

In the top-down effect study, Lawrence et al. 2018 had participants remember an orientation in a visual working memory task. The laminar activation of V1 was measured during the working memory delay period, when no orientation was present on the screen yet the participant held it in memory. Lawrence et al. 2018 found that the maintenance of visual working memory information activated the superficial and deep layers of V1 more

than the middle layers. This indicated that visual working memory information arrived top-down from higher-sensory areas into V1.

Another study by Lawrence et al. 2019 investigated whether laminar activation of V1 was distinct for feature-based attentional (top-down manipulation) and stimulus-contrast based (bottom-up manipulation) signals. In this task participants had to attend to a plaid stimulus which concurrently had two orientations present. The stimulus-contrast of the plaid was either high or low contrast on a given trial. After subtracting the activity of the high-contrast by the low-contrast stimuli for the bottom-up manipulation, Lawrence et al. 2019 found an increase in activity within the middle layers of V1 compared to the superficial and deep layers. This provided evidence that higher activity within V1 for higher contrast stimuli is partially explained by a higher activation of the bottom-up layers.

The current replication study followed the same behavioral paradigm and a similar fMRI design as Lawrence et al. 2018. To replicate the top-down laminar activation, the current study used the same visual working memory retro-cue task as selected by Lawrence et al. 2018. Additionally to the working memory task, Lawrence et al. 2018 had participants view an on-off subsequent presentation of two orientation stimuli. This task was used to localize orientation-selective voxels in V1, populations of neurons which activate more for one given orientation compared to another. Localizing these voxels was needed for the analysis of the working memory task. To concurrently aim to replicate the bottom-up laminar profile, the present study's localizer stimuli consisted of high and low contrasts, just like in the plaid stimuli of Lawrence et al. 2019.

To our surprise, the results from Lawrence et al. 2018 and Lawrence et al. 2019 did not replicate at 3T. The study found a significantly higher difference in data quality when comparing a previous 7T sample to the current 3T dataset. The bottom-up stimulus-contrast laminar effect showed no significant difference between agranular and granular depths, while the top-down visual working memory main effect could not be replicated, due to weakly distinguishable orientation-selective voxel populations at 3T. These results provide evidence against laminar fMRI at 3T using the current sequence and analysis pipeline.

Methods

Participants

Twenty-five right-handed healthy participants (13 females, mean age of 24, age range 18-46) with normal or corrected vision participated in both the contrast and working memory experiments of the study. Each participant gave written consent and filled out a behavioral and fMRI screening form. The local ethics committee (CMO region Arnhem-Nijmegen, The Netherlands) approved the study. Participants (except for two – researcher JKD and a PhD student within the FPdL lab) received a 8€ per hour compensation for the psychophysics part and a 10€ per hour compensation for the fMRI session. On average, participants earned 28€, as most participants completed a 1-h psychophysics session and a 2-h fMRI session. Data from two participants was subsequently removed due to excessive movement or not following task instructions. 23 participants were used for the bottom-up stimulus-contrast analysis. Another two participants either did not complete any working memory runs or their working memory data could not be realigned. 21 participants were used to analyze the top-down working memory data.

Another seventeen participants were taken from Lawrence et al. 2018 for the 7T comparative data quality analysis. This was not the whole dataset as some data within the repository were missing for the other participants.

Sample Size Estimation

The main aim of this study was to replicate previous results from Lawrence et al. 2018 and Lawrence et al. 2019. Since we used a lower static-field magnet to replicate data from a higher static-field, the loss in SNR was supposed to be compensated by the sample size. In the preregistration (Degutis et al. Jan 9, 2020), we calculated that a projected sample size of 60 participants would achieve 80% power to detect a between-subject effect size of 0.8. The calculation used both between (0.00436) and within-subject variances (0.427) of the laminar working memory effect (agranular layers subtracted by granular layers) from Lawrence et al. 2018 (Baker, Vilidaite, Lygo, … & Andrews, 2019). To compensate for the increased noise when scanning at (approximately) half the static magnetic field (3T compared to 7T), the between-subject variance from Lawrence et al. 2018 was multiplied by a factor of two.

However, due to the COVID-19 pandemic, data from only 25 out of the initially planned 60 participants was collected.

Tasks

Working Memory Task

The visual working memory retro-cue task was completed by participants during the psychophysics session and was the first task during the fMRI session (Figure 1a). In the beginning of a trial, the participants were presented with two sinusoidal Gabor gratings (following the parameters of Lawrence et al 2018: spatial frequency of 1 cycle per degree and a random phase, 50% contrast, edge softening linear ramp 0.5mm from the edge) of different orientations. The presented gratings (either 45° ± 3 or 135° ± 3°) were almost orthogonal to one another on each trial and were presented centrally. The sequence of presentation varied from trial to trial. Each grating was presented for 0.2s with the inter-stimulus interval being 0.4s. The onset of the grating was synchronized to the MRI. The two gratings were followed by a cue (presented for 0.8s) indicative of which stimulus to remember (i.e. maintain in working memory). The participant had to remember the orientation throughout a delay period of 10.7s. Following the working memory delay period, a probe stimulus was presented for 0.5s; the probe was of the same orientation as the remembered grating but slightly shifted clockwise or counterclockwise. The participant had 2 seconds to respond which way the probe was shifted compared to the remembered orientation. Participants had to respond with a (fMRI compatible) button-box; the left button (right hand's index finger) corresponded to counterclockwise and the right button (right hand's middle finger) corresponded to a clockwise shift. A feedback screen was presented for 0.3s indicating whether the answer was correct or incorrect. If a participant did not respond during a trial, the trial was marked as incorrect. An inter-trial-interval of 10s followed the feedback screen. Throughout the whole trial, a fixation cross was presented (a visual angle of 0.25° and 0.05° for the outer and inner circle, respectively; Lawrence et al. 2018 had a fixation dot). The participants were instructed to fixate on this cross. Each trial was 25.6s long. One run lasted for 521.6s.

The purpose of the participant's response was to ensure that they were performing the working memory task. After each trial the participant's response (clockwise or counterclockwise) was recorded and compared to the true orientation shift. This information was used by a QUEST staircase which calibrated their performance to 75% accuracy, optimizing for participant engagement in the task. After each trial, the degree difference between the remembered orientation and the probe was adjusted based on participant's performance. For example, the smaller the difference between the presented probe and the remembered orientation – the more difficult the trial.

As previously mentioned, participants engaged in the working memory task in both the scanner and psychophysics sessions. The aim of the latter was to determine the working memory accuracy of each participant prior to them entering the scanner. The psychophysics session occurred a day of two before the fMRI scan. A participant completed 120 trials of the working memory task. The stimuli were presented on a 24 inch BenQ XL2420T monitor (resolution 1920 x 1080, refresh rate 120 Hz) and the participants were seated 60 cm away from the screen; same monitor and distance as in Lawrence et al. 2018.

The working memory accuracy was measured as the orientation difference between the probe and the remembered stimulus. This number and an additional 25% increment (e.g. 5°+1.25°) was used as an initial value for the QUEST staircase within the scanner. The increment was added to account for a possible initial drop in performance due to the change in environment from a behavioral cubicle to an MRI scanner.

Contrast Task

Unlike the working memory task, the stimulus-contrast task was only done during the fMRI session (Figure 1b). Following Lawrence et al. 2018 localizer task, participants viewed either a clockwise of counterclockwise grating stimulus presented on-off at 2Hz (250ms on and 250ms off) in a single block of the contrast task. The stimulus parameters corresponded to the working memory task except for the stimulus contrast. On a single block, the stimulus contrast could have either been 30% (low) or 80% (high) (in Lawrence et al. 2018 only high stimulus-contrast gratings were presented). The grating stimulus was presented for 12.8 seconds. A baseline period with a blank screen followed for also 12.8 seconds. The orientation altered from block to block, while the stimulus contrast altered every two

blocks; meaning, both orientations were presented at one given contrast before moving onto the next contrast. The initial contrast was determined randomly for each run of each participant, while the orientation always started as 45°. The phase of the grating stimulus changed on each new presentation. The initial phase of the presented grating was determined randomly for each block.

As for the task, the participant had to fixate on a centrally presented fixation cross (same parameters as in the working memory task) and had to indicate when the fixation cross as a whole changed luminance (i.e. flashed). One flash took 250ms, or the same time as one on-off flip of the grating. In one block there could have been anywhere from one to four flashes (mean, median and mode being two flashes per block). The participant had to use their right hand's index finger to press a button using a fMRI compatible button box. One contrast run consisted of 32 contrast blocks (16 for each orientation). The total length of one run was 828.8s. Most participants completed two blocks.

Both the working memory and contrast tasks were programmed using Psychtoolbox (Brainard, 1997) with MATLAB version 2016b (MathWorks, Natrick, MA). Code developed by Lawrence et al. 2018 was used and adjusted for these two tasks.

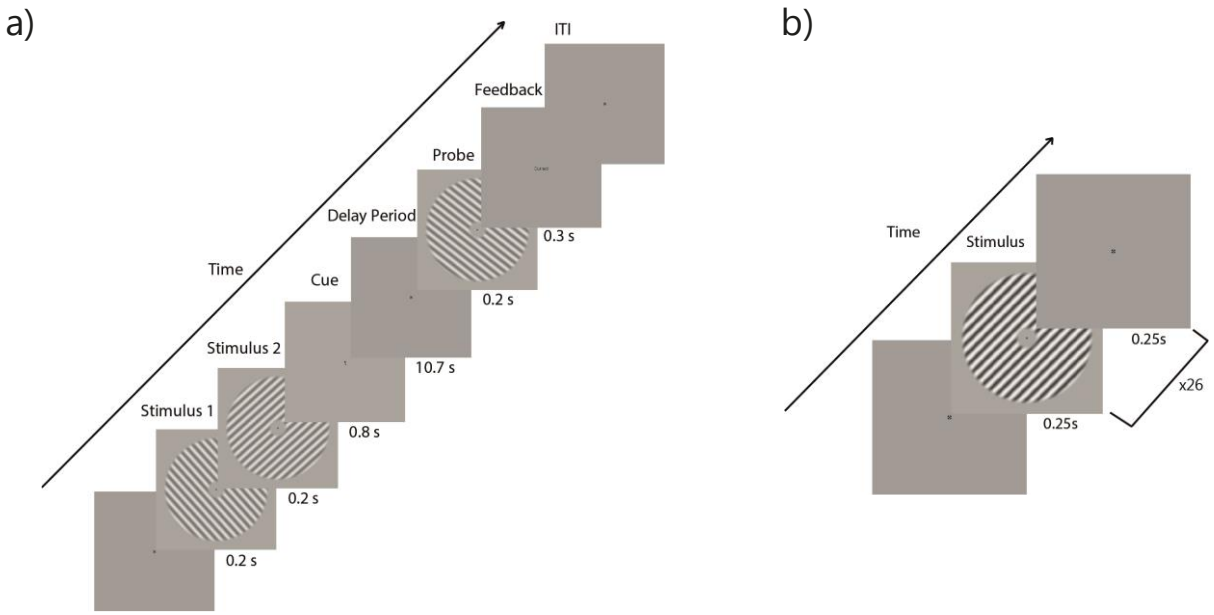


Figure 1: Stimulus presentation for the visual working memory and contrast tasks

(description continues on next page)

(a) On each trial a participant was presented with two orientation gratings and had to remember one of them based on the retro-cue provided after stimulus presentation. A 10.7s delay period followed during which the participant had to maintain the stimulus orientation in working memory. After the delay, the participant was presented with a probe and was asked to indicate whether the probe was shifted clockwise or counterclockwise compared to the remembered orientation. Feedback was provided after a judgment was made.

(b) On each contrast block a participant was presented an on-and-off flashing orientation grating for 12.8s. A blank screen followed for another 12.8s after the presentation. The participant's task was to indicate when the fixation cross changed luminance. The stimuli could have either been of low (30%) or high (80%) contrasts.

fMRI Data

Acquisition

FMRI data were acquired using 3T PrismaFit and Prisma Siemens scanners (Siemens, Erlangen, Germany) using 32 channel RF coils. Functional data were acquired using a T2*-weighted 0.8 mm isotropic voxel 3D gradient echo EPI sequence with a 4x2 acceleration and partial-Fourier reconstruction (3.2s TR, 31ms TE, 56 slices, 20° flip angle, 204.8x204.8x44.8 mm FOV). As in Lawrence et al. 2018 and 2019, the functional slices were manually adjusted for each participant by centering the slab on the Calcarine sulcus, thus making sure the early visual cortices were included in the FOV. A T1-weighted MPRAGE sequence (0.8mm isotropic voxels, 2.2s TR, 26.4ms TE) was used to acquire a high resolution anatomical image. A BOLDScreen (32 inch screen, 1920 x 1080 full color IPS LCD, 120Hz panel drive, https://www.crsltd.com/tools-for-functional-imaging/mr-safe-displays/boldscreen-32-lcd-for-fmri/) was used to display the stimuli within the MRI.

Data Preprocessing

Anatomical and functional scans were cropped to only include the occipital cortex and to remove the frontal cortex. This step reduced the brain area for which precise motion alignment and coregistration was needed. The preprocessing procedure was initially

separated for anatomical and functional scans and followed the preregistration (Degutis et al., Jan 9 2020)

Freesurfer's (http://surfer.nmr.mgh.harvard.edu/) recon-all function was used to segment the T1-weigted anatomical scan into three surfaces: CSF, gray matter and white matter.

SPM12 was used to motion correct functional scans (the parameters used were the same as in Lawrence et al. 2018: quality 1, interpolation 4-6, FWHM 1mm, sep 1, RTM 1). After each realignment, all scans were manually checked for remaining movement artifacts. Some functional scans could not be motion corrected using the SPM automated script, thus, due to excessive movement of seven participants, realignment was done by first realigning each functional run independently, followed by the realignment of the first volume of the first run to the first volumes of all other runs. These volume-to-volume transformation matrices (from run N to run one) were then used for all other volumes of the Nth run. This type of procedure reduced the number of possible local minima for the SPM algorithm to falsely terminate at.

After realignment, the functional scans were bias corrected using FAST (the same parameters as in Lawrence et al. 2018 were used: fast -t 2 -n 2 -H 0.1 -I 4 -l 20.0 --nopve -B). Functionals scans were also high-pass filtered with the OpenFmriAnalysis toolbox (van Mourik, 2019) using a cut-off of 52 s (two times the length of a single trial or block), thus removing low frequency scanner drift. Lawrence et al. 2018 had a more stringent high-pass filter of 28s. This seemed too stringent as it might have filtered-out actual signal, thus we decided to adjust the cut-off.

Following the coregistration procedure indicated in the preregistration, three steps were taken to align the anatomical image surfaces to the functional image. First, a simple rigid body transformation within the OpenFmriAnalysis toolbox (van Mourik, 2019) aligned all cortical surfaces to the mean functional image. In the case of the seven manually realigned participants, the mean image was replaced by the first volume of the first run. Second, OpenFmriAnalysis toolbox's boundary-based regression (BBR) was used for the registration of the mean image with the three anatomical volumes. Third, OpenFmriAnalysis was used to perform multiple recursive iterations of BBR (optimization method: Greve Fischl, accuracy:

30, number of iterations: 6) on the initial BBR output from step-two. The third step followed recent evidence from van Mourik et al. (2019) which found that recursive boundary registration improved coregistration in high-resolution images.

Functional Masks

In accordance to the preregistration (Degutis et al., Jan 9 2020), only the primary visual cortex was considered as a region of interest (ROI) in this study. Lawrence et al. 2018 had looked at the laminar profiles of all areas of the early visual cortex (V1-V3) and had only found significant results for V1. Furthermore, V1 was defined for each participant individually using a probabilistic anatomical atlas (Neuropythy package) instead of using retinotopic mapping as in Lawerence et al. 2018.

Following the preregistered analysis (Degutis et al. Jan 9, 2020) and Lawrence et al. 2018, two orientation-selective masks had to be defined in order to analyze the top-down working memory effect. The stimulus-contrast runs were used as a localizer to determine orientation-selective voxels for clockwise (45°) and counterclockwise (135°) orientations. To determine these functional masks, regressors of each orientation were contrasted against one another; only the 80% stimulus-contrast trials were used for the top-down working memory analysis. This provided two t-contrasts: one for the clockwise and one for the counterclockwise preferring voxels (the largest positive t-values for one contrast were the largest negative values for the other contrast). To determine the final orientation-selective functional masks, 500 voxels having the highest t-values for each t-contrast were chosen. This resulted in two functional masks: top 500 voxels for clockwise>counterclockwise and top 500 voxels for clockwise<counterclockwise.

Additionally, another functional mask was determined by contrasting all activity (orientation and contrast regressors) against baseline while using both 30% and 80% stimulus-contrast trials. This provided a t-contrast of the highest activated voxels across all conditions within V1. The top 500 voxels were used from the all>baseline t-contrast to define the mask.

Following Lawrence et al. 2018 and the preregistration (Degutis et al., Jan 9 2020), the t-values from the orientation-selective functional masks were used to weight the timecourses of each run. For each run, the z-scored BOLD signal within a voxel was multiplied by its corresponding t-value in a given mask. For example, the clockwise orientation mask's t-

values were multiplied by voxels' BOLD signal to give clockwise-weighted functional scans. This was done for the clockwise and counterclockwise masks on all functional runs. This weighting procedure was not applied to the activity vs. baseline mask, as it would have used t-values determined from the localizer to weight the localizer BOLD data (i.e. double-dipping).

An additional cross-validation control analysis was run to determine whether the orientation-selective masks had consistent t-values across trials. The analysis was run for each participant in both the 3T and 7T datasets for the high (80%) stimulus-contrast trials. On a given iteration, clockwise and counterclockwise regressors in two trials (two trials for each orientation) were assigned a different regressor name. For the sake of clarity, we referred to them as clockwise test and counterclockwise test, while the original orientation regressors were named clockwise train and counterclockwise train. A pairwise contrast was applied to both the test and train regressors resulting in four masks per iteration: clockwise>counterclockwise test mask, clockwise>counterclockwise train mask, and the two inverse contrast masks. The top 500 voxels with the highest t-value within V1 were taken from both the clockwise and counterclockwise train masks. The 500 train mask voxels were used to select the t-values of the corresponding voxels within the clockwise and counterclockwise test masks. These 500 t-values for each mask were then averaged. The logic being: if there was orientation-selective, then the test t-values for a given orientation should be above zero. In total, there were eight cross-validation iterations per contrast run per participant.

Measurement of Laminar Effect

Laminar Activity Estimation

In accordance to the preregistration (Degutis et al., Jan 9 2020), depth-specific activity was determined by applying a spatial GLM (van Mourik et al. 2018) to the functional data using the OpenFmriAnalysis toolbox (van Mourik 2019). Freesurfer's anatomical volumes were assigned into five equivolume depth bins consisting of one CSF bin, three gray matter bins, and one white matter bin. Since the anatomical surfaces had been previously coregistered to the mean functional image, the contribution of these five depth bins was determined for each voxel. For example, one voxel could have consisted of CSF, the first gray matter depth,

and the second gray matter depth in the same proportion; while another voxel could have been primarily white matter. Thus, a matrix of weights was generated for all voxels and the five depth bins indicating each voxel's proportional contribution to the respective depths. This matrix was regressed against the functional data resulting in one timecourse per depth bin per functional mask per functional run. Thus, in total there were five timecourses for all functional masks and functional runs (weighted by separate t-contrasts); however, only the three gray matter timecourses were used for the functional analyses.

Working Memory Activity Measurement

The working memory effect was measured using an approach described by Albers et al. (2018), outlined in the preregistration (Degutis et al., Jan 9, 2020), and used by Lawrence et al (2018). During a given trial, the remembered orientation was deemed as the preferred orientation, while the non-remembered one was non-preferred (e.g. 45° is preferred if the participant remembered the clockwise orientation). Since there were only two possible orientations that a participant could remember throughout the whole experiment, each trial consisted of a preferred and nonpreferred orientation and timecourse. In total there were an equal number of both clockwise and counterclockwise orientations as the preferred. Both the preferred and non-preferred timecourses were shifted by subtracting the average baseline across all five layers of the 0^{th} TR (the TR at 0s).

The activity within the two TRs (TR number 4 and 5) was used to estimate the working memory effect. These TRs were chosen since they occurred during the working memory delay period during which the perceptual activation from the initially presented grating stimuli had decreased, while the probe stimulus had not yet been presented. Lawrence et al. (2018) used TRs 3-4; however they also had a longer TR length, thus the signal from perceptual activation had already decreased. The activity was averaged across the two TRs within both preferred and non-preferred trials. To determine the main working memory effect, in each trial the activity of the preferred orientation voxels (the orientation that was remembered) was subtracted by the non-preferred orientation voxels (the orientation that was not remembered).

Contrast Activity Measurement

As outlined in the preregistration (Degutis et al., Jan 9 2020) the contrast task was used to determine the bottom-up laminar activity within V1. The activity vs. baseline mask was used to estimate the contrast effect. The trials were subdivided into high-contrast and low-contrast timecourses. As in the working memory analysis, both timecourses were shifted by subtracting the average baseline across all five layers of the 0^{th} TR (the TR at 0s). The activity within three TRs (3rd-5th) of the plateaued BOLD signal was used to calculate the bottom-up activity. The activity of the three chosen TRs was averaged for both high-contrast and low-contrast trials. For the stimulus-contrast effect the average high-contrast response was subtracted by the low-contrast averaged activity for each layer.

Bayesian Analysis

In addition to frequentist statistics and as indicated in the preregistration (Degutis et al., Jan 9 2020) we performed Bayesian analyses on the laminar effects of the top-down working memory and bottom-up stimulus-contrast tasks. MATLAB's bayesFactor toolbox (Krekelberg, 2020) was used to perform Bayesian one-sample t-tests using the default parameters (Cauchy prior of 0.707). Lee and Wagenmakers' (2014) criteria were used to interpret the Bayes Factors.

Results

Data Quality Comparison between 3T and 7T

Before evaluating the bottom-up and top-down cognitive effects in V1, we compared the data quality between the 3T and 7T datasets. Raw BOLD functional scan horizontal slices from a 3T and 7T participant are seen in Figure 2a) (bottom and top, respectively; see Supplementary Figure 1 for enlarged images). The 7T image shows a higher contrast between gray and white matter, while no such contrast can be seen in the 3T image. Furthermore, the 3T volume has higher white noise, especially towards the upper part of the slab (middle of the brain).

Temporal SNR (tSNR) was used in order to compare the noise within the datasets. The mean and standard deviation across time was calculated for the localizer runs (one localizer run for the 7T data and two localizer runs for 3T) for 17 participants from Lawrence et al. (2019) and the 23 participants from this study. To quantify the tSNR, the mean was divided by the standard deviation for each run. Two horizontal slices of the tSNR from two participants from 3T and 7T can be seen in Figure 2b) (bottom and top, respectively; see Supplementary Figure 1 for enlarged images). The 7T slice has overall higher tSNR (the scales are the same across both slices). The 3T slice has higher tSNR around the edges of the slab and lower (darker) tSNR at the top of the slab. This corresponds to the increased white noise seen in the raw BOLD image. We calculated two average values per participant which included the tSNR of the slab (only the occipital lobe) and V1. These values were averaged across participants (results seen in Figure 2c). The tSNR of 7T was greater both in the whole slab ($p=3.32e-06$, $t=5.44$, $df=38$) and in V1 ($p=9.01e-04$, $t=3.60$, $df=38$) when compared to the 3T data.

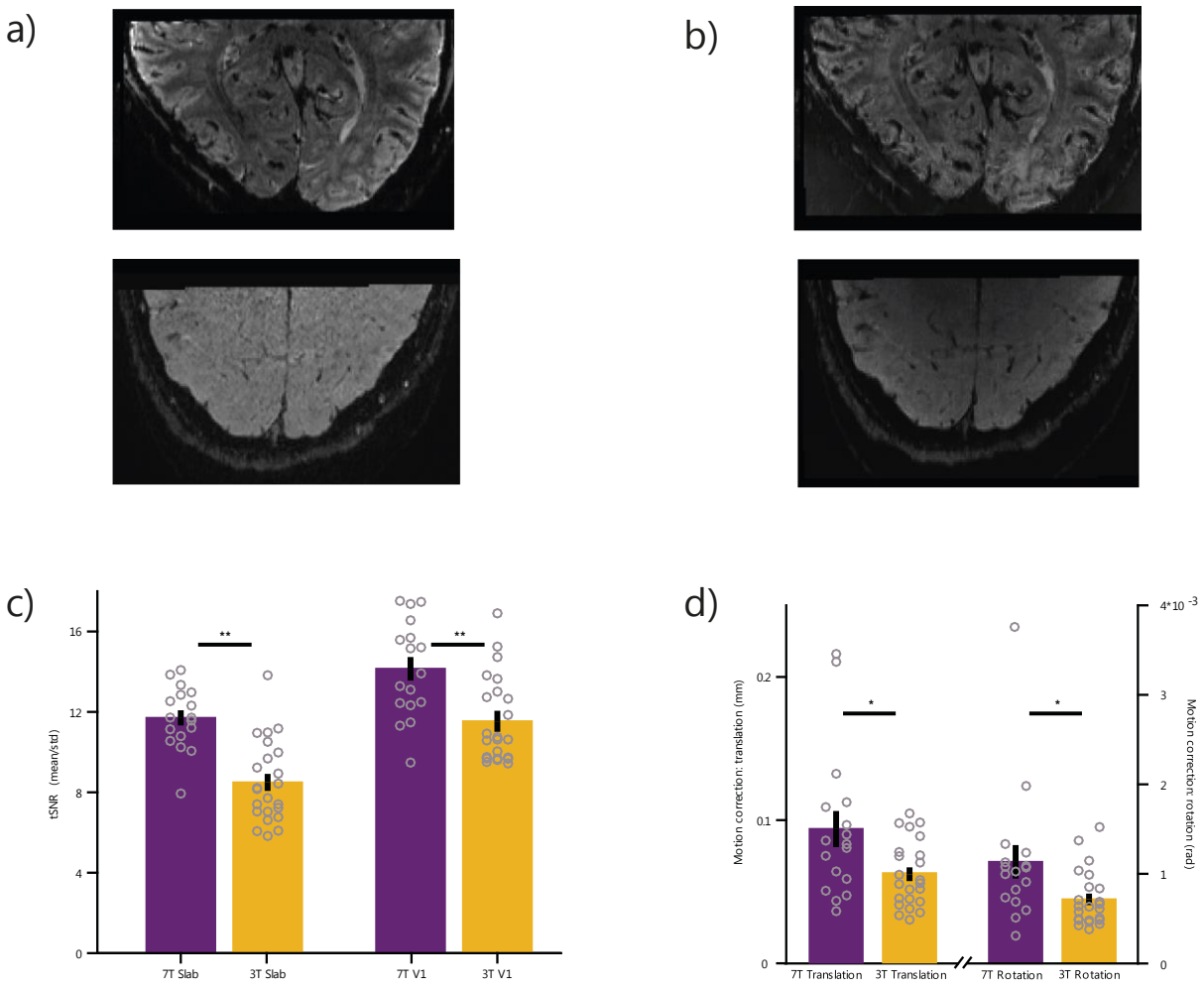


Figure 2: Data quality measurements comparing 7T and 3T functional scans

(a) The motion corrected raw BOLD functional data from a horizontal slice of a 7T participant (top) and a 3T participant (bottom). (b) The temporal SNR (temporal mean divided by temporal standard deviation) of BOLD functional data from a horizontal slice of a 7T participant (top) and a 3T participant (bottom). The scale for both tSNR images ranges from 0-25 (c) Quantified tSNR data across all contrast/localizer runs and participants from 7T (purple) and 3T (gold). The error bars show between-subject standard error while the gray circles indicate each participant's result. Temporal SNR was calculated for the whole slab (only the occipital lobe) and V1 (defined by retinotopy for the 7T participant and a probabilistic anatomical atlas for the 3T participant). (d) Average translation and rotation motion correction parameters for 7T (purple) and 3T (gold) participants. The left x-axis displays average translation (mm), while the right x-axis corresponds to average rotation (radians) of each participant. Motion parameters were averaged by the square-root of sum

of squares for translation and rotation independently. Asterisks denote significant two-sample t-tests (p<0.05, **p<0.005).*

In addition to tSNR, we also looked at the motion correction values from the 3T and 7T data. We compared the amount of translation and rotation. Following previous motion summation methods for the three translation values and three rotation values, we calculated the motion by summing the squared per-volume difference of the motion parameters, taking the square root, and averaging across the number of participants (Kennedy & Courchesne, 2008). The results showed higher motion for both translation (p=0.014, t=2.59, df=38) and rotation (p=0.023, t=2.36, df=38) for the 7T data.

Orientation Selectivity

In order to evaluate the top-down working memory effect, we had to make sure that the localizer voxels had sufficient signal to distinguish orientation-selectivity within V1. Thus, we used the localizer data to compare orientation-selectivity of voxels within V1 between 3T and 7T data. A linear SVM classifier (MATLAB's fitclinear function) was used to decode the presented orientation in a given localizer trial. Only the high-contrast trials from the 3T dataset were used to directly compare to the 7T dataset, as the latter experiment's localizer trials only consisted of high-contrast stimuli. For each trial in both datasets, the z-scored (across time for each run) values of thee TRs where the BOLD signal had plateaued were averaged to get one datapoint per voxel per trial. The classifier was trained in a leave-two-trials-out manner (one of each orientation) and cross-validated across all trials. The classifier accuracy across all validations for a given participant was averaged. The individual classifier accuracies, shown as purple or yellow dots, indicate above-chance decoding for each participant both in the 7T and 3T datasets (Figure 3a). The boxplots depict the bootstrapped individual null distributions calculated by permuting the classes 1000 times. The averaged results showed an increased classification accuracy for the 7T results as compared to the 3T (p=1.60e-03, t=3.39, df=38).

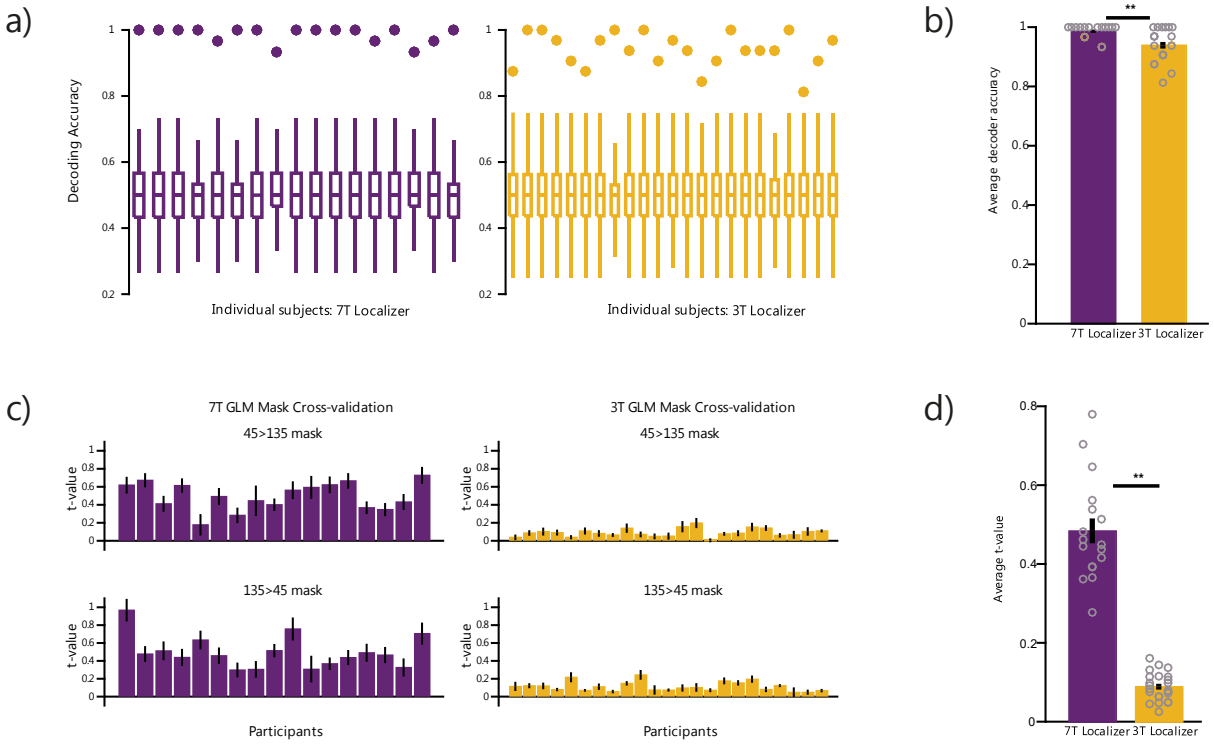


Figure 3: Orientation selectivity across 7T and 3T datasets

(a) Individual participant orientation decoding accuracy scores in localizer/contrast trials of 7T (purple) and 3T (gold) datasets. The filled-circles denote the accuracy score for a given participant while the box plots show bootstrapped individual null distributions. The decoding accuracy was calculated by dividing the number of correctly guessed orientations by the number of guesses across all iterations. There were two classes in total (clockwise and counterclockwise). The null distribution was calculated by randomly permuting the classes for each participant 1000 times and calculating the decoding accuracy score for each permutation. The error bars in the box plot denote the maximum and minimum value, while the top and bottom of the 'box' show the first and third quartile, respectively, while the middle stroke in the 'box' denotes the median. Only the high-contrast 3T trials were used for the decoding. (b) Average decoding accuracy across all 7T (purple) and 3T (gold) participants. The error bars show the between-subject standard error. The gray circles denote each participant's individual decoding score. (c) Individual participant top 500 voxel GLM mask cross-validation analysis of 7T (purple) and 3T (gold) participants. The bar graphs denote the average t-value across all iterations, while the error bar is the standard error across the iterations. The top two graphs show the clockwise (45°) preferring mask, while the bottom two graphs show the counterclockwise (135°) preferring mask. (d) Average GLM

*cross-validation t-value across both masks and participants. Error bars depict the between-subject standard error. The gray circles show each participant's individual average t-value across both masks. Asterisks denote significant two-sample t-tests (**p<0.005).*

In addition to the classification approach, we also cross-validated the GLM masks in both the 3T and 7T data. This analysis aimed to determine whether voxels within the masks were selective to either the clockwise or counter-clockwise orientations (see methods section for more information). Each participant provided two data points: one for the clockwise and one for the counter-clockwise orientation voxel masks. A positive t-value indicated selectivity to the given orientation, while a negative t-value indicated a higher on-average activation for the opposite orientation. The two values for the participant were averaged, thus providing one cross-validated contrast value per participant. The masks used were created on the 3T and 7T localizer data sets. As in the classifier cross-validation analysis, the 3T data only used the high-contrast stimuli for the mask cross-classification. Individual results showed all positive t-values for participants in the 7T dataset for both orientation masks. Most 3T dataset t-values for both orientation masks were positive; however, they had lower values compared to the 7T results. The results averaged across all participants of the two datasets showed a higher t-value activation for 7T data compared to the 3T data (p=1.56e-16, t=13.9, df=38).

The high-contrast 3T orientation masks were used to measure the working memory laminar effect, while the activity vs. baseline mask used for the contrast analysis was based on both high and low stimulus-contrast localizer runs. The orientation mask using both high and low contrast localizer stimuli did not have orientation selectivity (one-sample t-test: p=0.563, t=0.587, df=22), thus the masks based on high-stimulus contrast were chosen. The contrast effect analysis did not require orientation-selective masks and using a high-contrast activity vs. baseline mask would have biased the main contrast effect by selecting for highly active high-contrast voxels.

Laminar Results for Bottom-Up Contrast Activity

Participants were presented with high and low stimulus-contrast clockwise and counter-clockwise orientations during the localizer task. The activity vs. baseline mask (i.e. 500 most active voxels) was used to measure the stimulus-contrast effect. The main effect of

stimulus-contrast was determined by subtracting the high stimulus-contrast activity by the low stimulus-contrast and averaging the three TRs where the BOLD signal had plateaued.

Figure 4 shows the results from the bottom-up contrast analysis. Figure 4a) shows the timecourse of high and low stimulus-contrast activity. The main stimulus contrast effect was highly significant (paired-sample t-test: p=1.27e-09, t=9.97, df=22) indicating a higher activation for high stimulus-contrast stimuli compared to low stimulus-contrast stimuli. In other words, the bottom-up manipulation of stimulus contrast worked. Figure 4b) shows the laminar effect of the stimulus-contrast; the layers have been separated into three bins: deep, middle, and superficial. A statistical analysis of the agranular (superficial and deep layer) activation compared to granular (middle layer) showed no statistical significance (one-sample t-test: p=0.778, t=-0.28, df=22). A Bayesian analysis provided moderate evidence (BF_{10}<1/3) towards the null (BF_{10} = 0.227).

Figure 4c) shows the distribution of the agranular vs. granular laminar effect across participants. It ranges from -0.06 (higher granular activity) to 0.04 (higher agranular activity). Figure 4d) presents a scatter plot of the main contrast effect and laminar effect. There is a trend towards a negative correlation (R=-0.535) between the main effect and the agranular vs. granular laminar effect (i.e. the higher the main stimulus-contrast effect, the higher the granular layer activation). However, it must be noted, that this correlation might be driven by the datapoint in the bottom-right of the scatter plot; no statistical test has been done on the correlation.

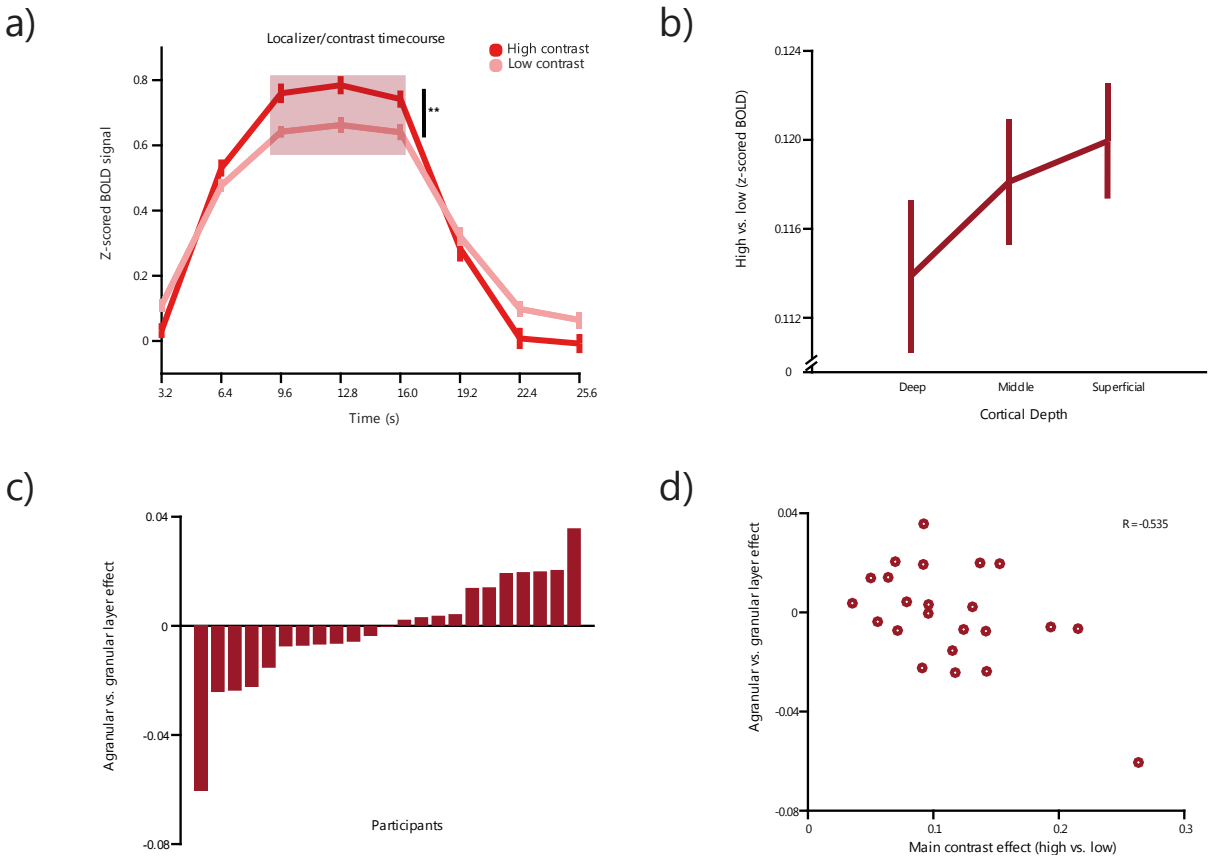


Figure 4: Stimulus-contrast bottom-up effects in V1 of the 3T data

*(a) Average timecourse across participants of high-contrast (red) and low-contrast (pink) stimuli. The error bars denote within-subject standard error. The stimulus was presented from the beginning of the trial to 12.8s; however, the main and laminar bottom-up effect was measured on the plateaued BOLD response in TRs at 9.6s, 12.8s and 16.0s (shown by transparent rectangle). Asterix denotes significant paired-sample t-test (**p<0.005). (b) Average laminar signal across cortical depth in V1. The laminar effect was calculated by subtracting the high-contrast by the low-contrast timecourse of the aforementioned TRs. The error bars depict within-subject standard error. (c) Agranular effect of each individual participant. The effect was calculated by averaging the agranular layer (deep and superficial depth) and subtracting the granular layers (middle depth). A negative agranular effect means that there was higher granular layer activation (middle depth), while a positive effect means that there was higher agranular layer activation (averaged deep and superficial). (d) Correlation between the main bottom-up effect and the laminar agranular bottom-up effect. Each red circle represents a participant.*

Laminar Results for the Top-Down Working Memory Activity

As the main task of the study, participants did a visual working memory experiment in which they had to remember either a clockwise (45°) or counter-clockwise (135°) orientation and maintain it throughout a delay period. At the end of a trial a new orientation was presented; the participant had to indicate whether the probed orientation was shifted clockwise or counterclockwise compared to the remembered orientation (mean threshold across participants: 6.87°, across participant threshold standard deviation: 3.91°, within-participant threshold average standard deviation: 0.129°). Orientation masks from the high stimulus-contrast trials were used for the analysis. On a given trial the remembered orientation, and thus the timecourse of that orientation's mask, was deemed the preferred orientation, while the non-remembered orientation was the non-preferred.

In accordance to previous univariate visual working memory analyses in the early visual cortex (Albers et al. 2018), the main working memory effect was calculated by subtracting the averaged preferred by the non-preferred timecourse; the two delay TRs (TR 4 and 5) were averaged together to get the main working memory effect (Figure 5a). Despite a small difference between the preferred and non-preferred orientations during the delay period TRs, there was no statistical difference between the two (paired-sample t-test: p=0.575, t=0.570, df=20). Figure 5b) shows the distribution of the main working memory effect. Despite more participants having higher activation for preferred compared to non-preferred (a positive working memory main effect), the majority of participants' main effect agglomerated around zero.

The lack of a main working memory effect made the interpretability of the laminar effect difficult. Yet following our preregistration, the laminar analysis was also conducted for the top-down working memory effect. Figure 5 c) shows a higher activation of the middle bin compared to the superficial and deep bins; however, the difference between agranular and granular activation was not significant (one-sample t-test: p=0.232, t=-1.23, df=20). A Bayesian analysis provided anecdotal (BF_{10}<0) evidence towards the null (BF_{10} = 0.443).

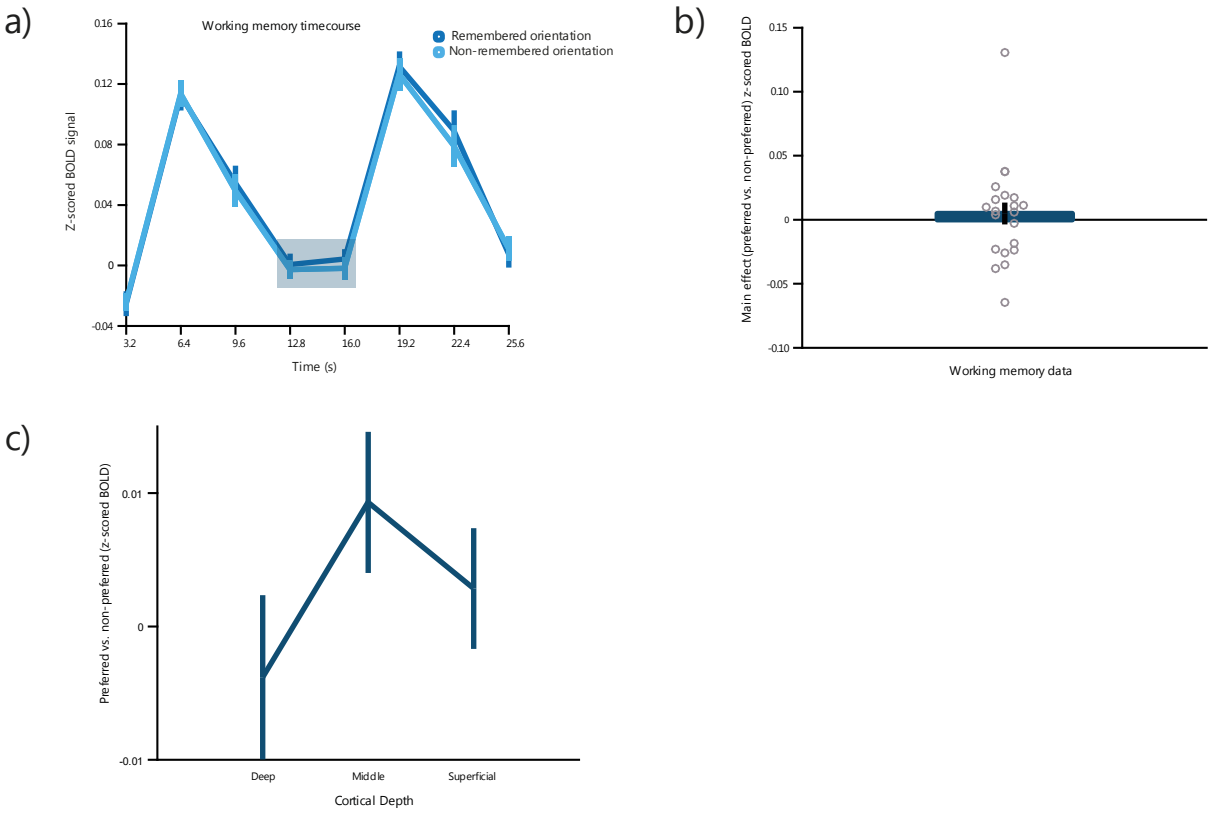


Figure 5: Visual working memory top-down effects in V1 of the 3T data

(a) Average timecourse across participants for remembered/preferred (deep blue) and non-remembered/non-preferred (light blue) stimuli. The error bars denote within-subject standard error. The delay period analyzed was TRs at 12.8s and 16.0s (shown by transparent rectangle). These were chosen since the BOLD response had decreased close to zero. (b) Main working memory effect depicting z-scored BOLD of preferred subtracted by non-preferred averaged across the delay period TRs. The error bar depicts within-subject standard error and the gray circles are individual participant results. (c) Laminar effect of the visual working memory delay period across cortical depth of V1. The error bars show within-subject standard error.

Discussion

The aim of the study was twofold: to provide a comparison of data quality when using laminar functional imaging at 3T and 7T, and to try to replicate previous 7T bottom-up and top-down laminar results at 3T. Participants in the study had to perform two tasks: a retro-cue working memory task and a stimulus-contrast task. The former task aimed to replicate the Lawrence et al. (2018) top-down working memory finding, while the latter tried to replicate Lawrence el al. (2019) bottom-up stimulus-contrast finding.

As expected, we found higher tSNR in the 7T functional scans in both the whole slab and V1. Furthermore, motion was higher within the 7T sample in both translation and rotation when compared to the 3T group. Orientation selectivity of the localizer scans was measured using a decoder and a cross-validation of GLM masks. The MVPA analysis showed above-chance decoding of orientation for all 3T and 7T participants, yet the 7T sample had higher decoding accuracy than 3T. Similarly, the cross-validation orientation mask analysis showed higher average t-values across iterations for 7T compared to 3T. Most individual participant orientation masks for 3T had positive t-values across trials (indicating orientation-selectivity); however, the values were lower compared to 7T. Finally, the stimulus-contrast bottom-up laminar analysis did not show any consistent laminar effect, while the top-down visual working memory results did not show a main effect during the working memory delay period, thus making the laminar analysis uninterpretable.

Data quality at 7T and 3T

The unsurprising difference in tSNR coincides with previous studies finding lower SNR at lower field strengths (Maubon, Ferru, Berger et al. 1999; Triantafyllou, Hoge, Krueger, et al. 2005). Yet, the result of higher motion correction within 7T compared to 3T constitutes a further explanation. One possibility is that higher motion correction at 7T might not have necessarily meant that there was higher motion in the 7T participant sample. Both studies used retrospective motion correction using the voxel-based registration (VBR) method employed in SPM's motion correction. The algorithm used a six-parameter rigid-body realignment based on a least-squares cost function which primarily relied on voxel intensity differences (SPM12). Since the noise within the 3T functional scans was high, as even seen from the images of the functional slices, slice-to-slice motion correction could have

prematurely stopped at a cost function's local minimum. This speculation follows the idea that the cost function state-space in the 3T functional images might have had shallower global minima. The lower tSNR might have resulted in lower voxel intensity differences, thus the algorithm might have struggled to determine the 'true' minimum. In contrast, a higher tSNR and contrast in the functional 7T scans would have resulted in larger differences between the global minimum and local minima, hence resulting in lower false realignment. This would result in under-corrected images and lower motion correction parameters at 3T, and would explain the higher motion parameters in 7T. This speculation is also corroborated by the inability to automatically realign seven participants at 3T (see Methods). The algorithm had to be 'helped' by first realigning the first run's first volume to the consecutive runs' first volumes and using these transformation matrices for the corresponding runs' other volumes. When the SPM algorithm initially failed, the lack of realignment was easily observed when examining the functional images. Since the above argumentation is speculation, another explanation – possibly more parsimonious – is that the 7T participants just moved more than the 3T participants. This explanation cannot be dismissed. A potential improvement for motion correction would the use of boundary-based regression (BBR). A new study showed better motion realignment for the high-resolution data using BBR compared to SPM's VBR (Huang, Carlin, Henson, & Correia, 2020). In our study BBR was only used to coregister the mean functional scan to the anatomical using the segmented anatomical.

Top-down working memory effect

The top-down visual working memory analysis relied on having orientation-selective masks within V1. Each trial had a preferred and non-preferred timecourse; if the participants remembered the clockwise (45°) orientation the clockwise orientation mask's timecourse for that trial was deemed as preferred, while the counterclockwise (135°) timecourse was non-preferred. Since there was the same number of trials where the remembered orientation was clockwise and counterclockwise, the preferred and non-preferred timecourses had the same number of each orientation's trials from their corresponding masks. The decoding and orientation mask cross-validation analyses compared the orientation selectivity within V1 between 3T and 7T high-contrast localizer trials in order to make sure that the working memory analysis could properly be done. The decoding results

indicated that orientation-selective information was present within both the 3T and 7T localizer runs. However, many participants showed ceiling effects as decoding accuracy was close to 1. Thus, the cross-validation of orientation-selective GLM masks was used as a more sensitive analysis, and we tried to see whether orientation-selective populations were consistent over trials. The large difference in average t-value across iterations between 7T and 3T indicated that the former's orientation-selective masks had a larger mean difference over time. The higher t-values at 7T might be explained by the lower signal in the 3T localizer.

We did not find higher preferred timecourse activation compared to non-preferred during the working memory delay period, while previous studies have shown this effect (Albers et al. 2018, Lawrence et al. 2018). This univariate difference is seen as an explanation for the many multivariate results for stimulus-specific working memory representations during the delay period in V1 (Harrison, & Tong, 2009; Sreenivasan, & D'Esposito, 2019) Hence, the main working memory effect is not in question; this failed replication was probably due to the lower average t-value in the cross-validation of GLM masks. More concretely, the 3T orientation masks must have had both 45° and 135° preferring voxels irrespective of the corresponding orientation-selective mask. This would have resulted in the average preferred timecourse being a mixture of voxels for the remembered and non-remembered orientations; likewise for the non-preferred timecourse. The size of the previous univariate working memory result at 3T (even though it was averaged across V1-V3 instead of just V1) was on par with the previous result at 7T (approx. 0.05 percent difference change; Albers et al. 2018; Lawrence et al. 2018). The increased spatial resolution at 7T counterbalanced the increased SNR which could have resulted in similar sizes of the effect. This begs the question whether the decreased tSNR in the high-resolution 3T functional scans in the current study might have resulted in the null finding. Even though V1 was active in the activity vs. baseline condition for all participants (activation was manually checked), the lower tSNR at 3T might not have enabled a distinction of 45° and 135° preferring voxels. The study could have been improved by having extra localizer runs, since in the current case the orientation-selective functional masks were calculated based on the same number of high-contrast trials at both 3T and 7T.

Bottom-up contrast effect

The bottom-up stimulus-contrast effect laminar results are harder to interpret, as they rely on both cognitive and data quality related arguments. Previous laminar fMRI research of V1 has shown a middle layer activation for stimulus-contrast (Lawrence et al. 2019; Marquardt, Schneider, Gulban, Ivanov, & Uludağ, 2019) and stimulus presentation (Koopmans et al. 2010); however a few laminar studies have not shown differential laminar activation for stimulus contrast (van Dijk, Fracasso, Petridou, & Dumoulin, 2020) or presentation (Lawrence et al 2018). In order to understand the possible reasons for a bottom-up laminar effect in some studies and a lack of one in others, the paradigms and analysis pipelines used should be evaluated. Van Dijk et al. (2020) had their participants fixate on flashing gratings of different contrasts which would shift after each flash. In separate runs, the participants had to either indicate when the direction of shift changed or had to passively fixate on the centrally presented fixation cross. Van Dijk et al. (2020) did not find any laminar specificity for the contrast modulation and only found a linear increase in BOLD activity across layers for increasing contrast. In addition to the working memory task, Lawrence et al. (2018) also analyzed the localizer which was operationalized as a bottom-up task. Similarly to the working memory analysis, they subtracted the presented (preferred) orientation by the non-presented (non-preferred) orientation timecourse. The laminar results did not indicate any specificity for the presented orientation's timecourse. Both van Dijk (2020) and Lawrence et al. (2018) results can be interpreted as a combination of both top-down and bottom-up signals within V1. Van Dijk et al. (2020) had both an attentional manipulation and a passive fixation task; in the latter the participants could have incidentally allocated their attentional resources to the grating irrespective of the task instructions. In Lawrence et al. (2018) the presented orientation was consciously perceived, thus both a bottom-up stimulation and a top-down global ignition (Dehaene & Changeux, 2011) could have occurred resulting in no differentiable laminar profile. More precisely, the presented (preferred) timecourse could have consisted of both top-down and bottom-up signals from the presented orientation while the non-presented (non-preferred) timecourse might have had remnants of the presented timecourse's top-down and bottom-up signals. Thus, a subtraction of presented by non-presented would have resulted in a negative linear shift of the combined top-down and bottom-up laminar effects.

In the studies where a middle-layer activation was found, the tasks or analyses isolated the bottom-up signals. Lawrence et al. 2019 had participants attend to either a clockwise or counterclockwise grating which were presented concurrently as a plaid stimulus. The stimulus-contrast varied across trials. In the analysis Lawrence et al. 2019 subtracted the attended high-contrast stimuli by the attended low; this might have also subtracted out the attentional effect (assuming it was the same across both high and low contrast stimuli) resulting in a higher middle layer laminar profile. Marquardt et al. (2019) had participants actively fixate on a centrally presented fixation cross and indicate when the cross's color changed by pressing a button corresponding to the particular color. In this attentionally intensive task, the participant's resources were allocated to the fixation cross instead of the grating stimulus. This does not necessarily mean that the participant did not consciously perceive the stimulus; however, the top-down signal might have been weaker compared to the bottom-up signal. Marquardt et al. also used a novel deconvolution analysis which aimed to remove the draining vein effect. This yielded higher activation within the middle layers compared to others.

Similarly to the localizer (bottom-up) task from Lawrence et al. 2018, in the current study participants also actively focused on the fixation cross and indicated when it changed luminance. The main difference between the two localizer tasks was the subtraction analysis. Unlike in Lawrence et al. 2018 where the presented stimulus was subtracted by the non-presented, the current study – following Lawrence et al. 2019 – subtracted the high-contrast stimuli by the low. This could have also subtracted the attentional top-down effect, assuming participants were as attentive to the high and low contrast stimuli. Thus, our prediction was that this task and analysis would have isolated the bottom-up signal and yielded a higher activation in the middle layers. However, the results indicated no consistency across participants for higher granular layer activation. The laminar results follow the draining vein effect consistently found in other gradient-echo BOLD laminar fMRI studies (Goense, Bohraus, & Logothetis, 2016; Markuerkiaga, Barth, & Norris, 2016; Uludağ, K., & Blinder, 2018; also see Supplementary Figure 2 for individual depth timecourses during the stimulus-contrast task which show higher superficial depth activation). If this result cannot be explained by cognitive effects, an explanation must depend on the data quality at 3T. Similarly to the null results from the top-down effect, the explanation for the bottom-up

null effect can be explained by lower tSNR at 3T. In this case, higher overall activation due to higher contrast stimuli might have resulted in a more pronounced draining-vein effect in the high stimulus-contrast trials compared to the low-contrast stimuli. This non-linear additive draining-vein effect would have not been canceled out when the low-contrast timecourse was subtracted. Thus, the current results might be interpreted as a combination of the draining-vein effect and bottom-up effect.

Feasibility of laminar fMRI at 3T

The current null results for the top-down and bottom-up tasks indicate that laminar fMRI at 3T is not feasible with the given sequence and analysis; however, previous studies at 3T have shown laminar-specific activation (Koopmans et al. 2010; Scheeringa et al. 2016). Scheeringa et al. 2016 found correlations between different oscillatory activity and corresponding layers of V1. They used a sequence with a higher repetition time (3.4s compared to our 3.2s) and fewer slices (48 compared to our 56), which would have allowed for higher signal per slice. Additionally, Scheeringa et al. interpolated the depth signals which resulted in 21 bins and averaged across the 10% most active voxels within a given region of interest. In contrast to this study's selection of voxels based on a GLM, averaging across a larger piece of cortex might have averaged out the noise. A new preprint has demonstrated that laminar fMRI is feasible both in 3T and in 1.5T but only when a larger portion of the cortex is averaged and the spatial resolution is reduced as a consequence (Markuerkiaga et al., 2020). Koopmans et al. 2010 used a sequence in which only 20 0.75mm isotropic slices were acquired in a volume TR of 60 seconds; this study was more of a proof-of-principle that differential laminar activity can be seen within V1; however, such a sequence is not feasible for most cognitive tasks due to the high acquisition time per volume. A sequence with slightly lower resolution and longer volume TR might have provided better results, as this could have increased the SNR.

Conclusion

The current study aimed to replicate previous 7T bottom-up and top-down laminar results using a 3T scanner and provide evidence for the feasibility of using a 3T scanner for laminar fMRI experiments. Previous results had found higher activation in V1's granular layers for the bottom-up effect and higher activation for V1's agranular layers for the top-down effect (Lawrence et al. 2018, 2019). We did not replicate these results: there was no differential laminar activation for the stimulus-contrast bottom-up task and the visual working memory top-down main effect could not be replicated making the laminar effect uninterpretable. These results can be partly explained by the following factors: reduced tSNR at 3T compared to 7T, a possible under-correction of motion at 3T, and the 3T dataset's lower t-contrast activation for orientation-selective voxel populations. A few previous 3T laminar fMRI studies have found laminar-specific activation (Koopmans et al. 2010, Scheeringa et al. 2016, Markuerkiaga et al. 2020); however, this study provides evidence against doing laminar fMRI at 3T.

References

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.

Albers, A. M., Meindertsma, T., Toni, I., & de Lange, F. P. (2018). Decoupling of BOLD amplitude and pattern classification of orientation-selective activity in human visual cortex. *NeuroImage, 180*, 31-40.

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2019). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *arXiv preprint arXiv:1902.06122*.

Balaram, P., Young, N. A., & Kaas, J. H. (2014). Histological features of layers and sublayers in cortical visual areas V1 and V2 of chimpanzees, macaque monkeys, and humans. *Eye and brain*, 6(Suppl 1), 5.

Barbas, H. (2015). General cortical and special prefrontal connections: principles from structure to function. *Annual review of neuroscience, 38*, 269-289.

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.

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

Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annual review of neuroscience*, *21*(1), 47-74.

Degutis, K. J., Ehinger, B. V., & de Lange, F. (2020, January 9). Laminar fMRI at 3T: A replication attempt of top-down and bottom-up laminar activity in the early visual cortex. Retrieved from osf.io/txuye

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*(2), 200-227.

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.

Goense, J., Bohraus, Y., & Logothetis, N. K. (2016). fMRI at high spatial resolution: implications for BOLD-models. *Frontiers in computational neuroscience*, *10*, 66.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632-635.

Huang, P., Carlin, J. D., Henson, R. N., & Correia, M. M. (2020). Improved motion correction of submillimetre 7T fMRI time series with boundary-based registration (BBR). *NeuroImage*, *210*, 116542.

Huber, L., Handwerker, D. A., Jangraw, D. C., Chen, G., Hall, A., Stüber, C., . . . Guidi, M. (2017). High-resolution CBV-fMRI allows mapping of laminar activity and connectivity of cortical input and output in human M1. *Neuron, 96*(6), 1253-1263. e1257.

Kennedy, D. P., & Courchesne, E. (2008). The intrinsic functional organization of the brain is altered in autism. *Neuroimage*, *39*(4), 1877-1885.

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.

Krekelberg, B. (2020). bayesFactor (https://github.com/klabhub/bayesFactor), GitHub. Retrieved July 29, 2020.

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, 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. e3434.

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, M., & Wagenmakers, E. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139087759

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.

Markuerkiaga, I., p Marques, J., Bains, L. J., & Norris, D. (2020). An in-vivo study of BOLD laminar responses as a function of echo time and static magnetic field strength. *bioRxiv*.

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.

Maubon, A. J., Ferru, J. M., Berger, V., Soulage, M. C., DeGraef, M., Aubas, P., ... & Rouanet, J. P. (1999). Effect of field strength on MR images: comparison of the same subject at 0.5, 1.0, and 1.5 T. *Radiographics*, *19*(4), 1057-1067.

Scheeringa, R., Koopmans, P. J., van Mourik, T., Jensen, O., & Norris, D. G. (2016). The relationship between oscillatory EEG activity and the laminar-specific BOLD signal. *Proceedings of the National Academy of Sciences*, *113*(24), 6761-6766.

Sharoh, D., Van Mourik, T., Bains, L. J., Segaert, K., Weber, K., Hagoort, P., & Norris, D. G. (2019). Laminar specific fMRI reveals directed interactions in distributed networks during

language processing. *Proceedings of the National Academy of Sciences*, *116*(42), 21185-21190.

Sreenivasan, K. K., & D'Esposito, M. (2019). The what, where and how of delay activity. *Nature Reviews Neuroscience*, *20*(8), 466-481.

Triantafyllou, C., Hoge, R. D., Krueger, G., Wiggins, C. J., Potthast, A., Wiggins, G. C., & Wald, L. L. (2005). Comparison of physiological noise at 1.5 T, 3 T and 7 T and optimization of fMRI acquisition parameters. Neuroimage, 26(1), 243-250.

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

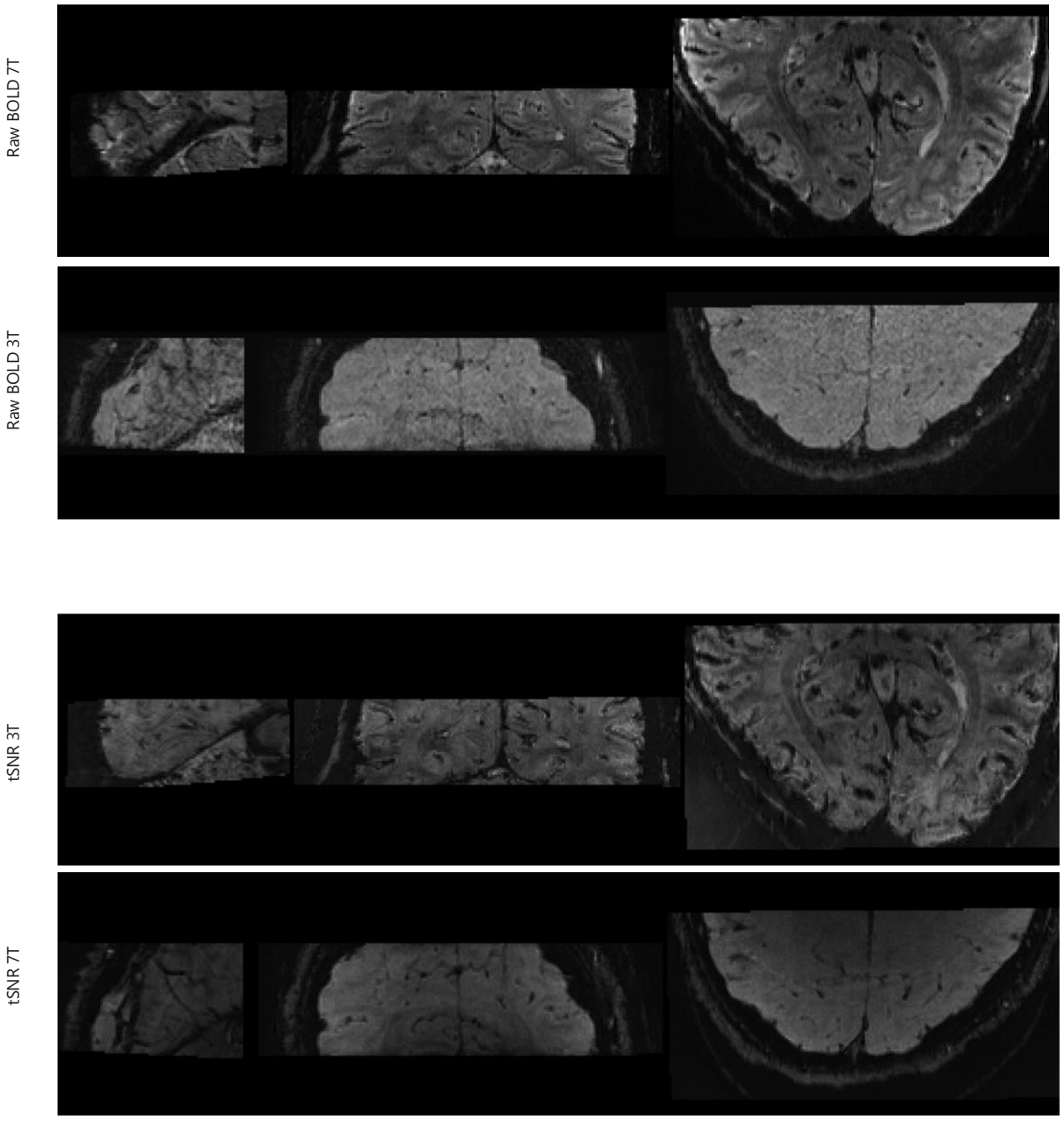
van Dijk, J. A., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2020). Linear systems analysis for laminar fMRi: evaluating BoLD amplitude scaling for luminance contrast manipulations. Scientific reports, 10(1), 1-15.

van Essen, D. C., Lewis, J. W., Drury, H. A., Hadjikhani, N., Tootell, R. B., Bakircioglu, M., & Miller, M. I. (2001). Mapping visual cortex in monkeys and humans using surface-based atlases. *Vision research, 41*(10-11), 1359-1378.

van Mourik, T., Koopmans, P. J., & Norris, D. G. (2019). Improved cortical boundary registration for locally distorted fMRI scans. *PloS one*, *14*(11).

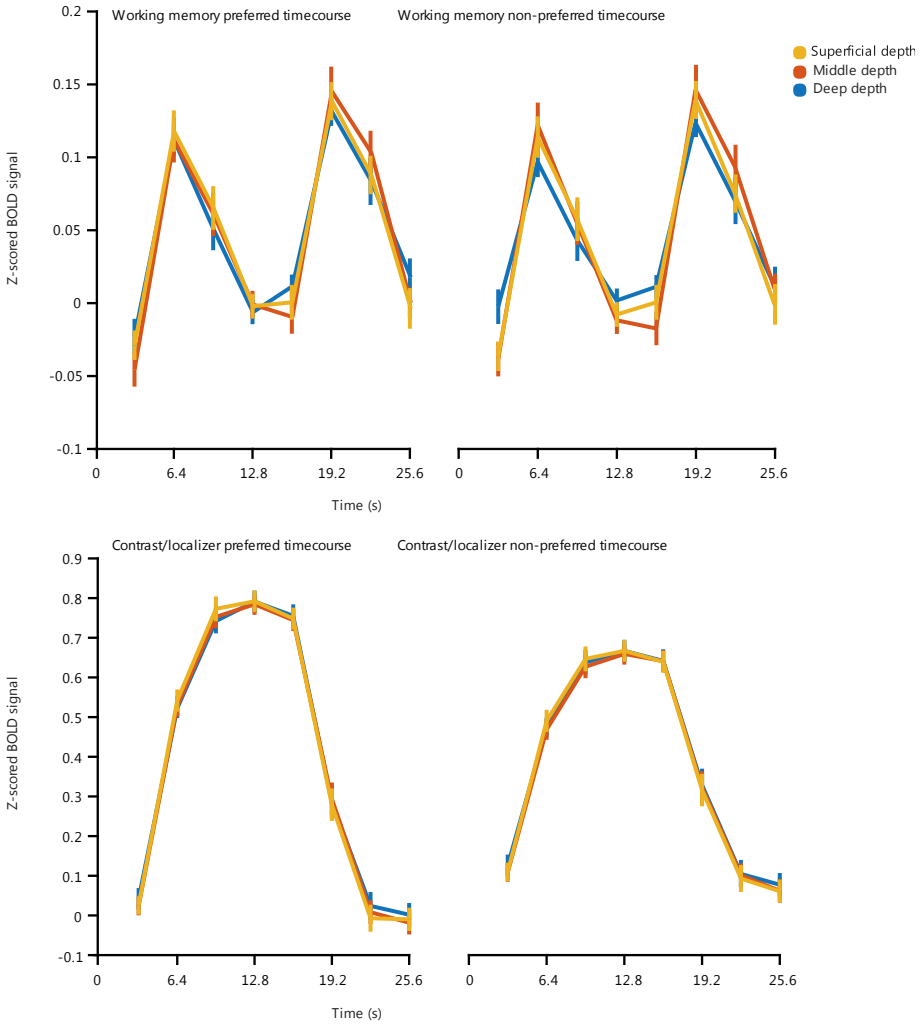
van Mourik, T., Van Der Eerden, J. P., Bazin, P.-L., & Norris, D. G. (2019). Laminar signal extraction over extended cortical areas by means of a spatial GLM. *PloS one*, *14*(3), e0212493.

Supplementary Figures



Supplementary figure 1: Qualitative data quality comparison between a sample 7T and 3T participant

The top two images show the motion corrected raw BOLD functional data from three slices of a 7T (top) and 3T (bottom) participant. The bottom two images show the temporal SNR of BOLD functional data from three slices of a 7T (top) and 3T (bottom) participant.



Supplementary figure 2: Individual cortical depth timecourses for the contrast/localizer and working memory tasks

The top two graphs show the activation of the three cortical depths during the working memory task for the preferred (left) and non-preferred (right) timecourses. The bottom two graphs show the preferred (left) and non-preferred (right) timecourses for the contrast/localizer task. Error bars show within-subject standard error.