Eye-selective fMRI activity in human primary visual cortex: Comparison between 3 T and 9.4 T, and effects across cortical depth

Natalia Zaretskaya\textsuperscript{a,b,c,d,*}, Jonas Bause\textsuperscript{d,e}, Jonathan R. Polimeni\textsuperscript{f,g,h}, Pablo R. Grassi\textsuperscript{c,d,i}, Klaus Scheffler\textsuperscript{d,e}, Andreas Bartels\textsuperscript{c,d,i}  

\begin{itemize}
  \item \textsuperscript{a} Institute of Psychology, University of Graz, Graz, Austria
  \item \textsuperscript{b} BioTechMed-Graz, Austria
  \item \textsuperscript{c} Vision and Cognition Lab, Centre for Integrative Neuroscience, University of Tübingen, Germany
  \item \textsuperscript{d} Max Planck Institute for Biological Cybernetics, Tübingen, Germany
  \item \textsuperscript{e} Department for Biomedical Magnetic Resonance Imaging, University Hospital Tübingen, Germany
  \item \textsuperscript{f} Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA
  \item \textsuperscript{g} Department of Radiology, Harvard Medical School, Boston, MA, USA
  \item \textsuperscript{h} Harvard-Massachusetts Institute of Technology, Division of Health Sciences and Technology, Cambridge, MA, USA
  \item \textsuperscript{i} Department of Psychology, University of Tübingen, Germany
\end{itemize}

ARTICLE INFO

Keywords:
- Ocular dominance
- fMRI
- 9.4 T
- V1
- Ultrahigh field MRI
- Cortical layers
- Cortical depth

ABSTRACT

The primary visual cortex of humans contains patches of neurons responding preferentially to stimulation of one eye (the ocular dominance columns). Multiple previous studies attempted to detect their activity using fMRI. The majority of these fMRI studies used magnetic field strengths of 4 T and higher. However, there have been reports of reliable eye-selective activations at 3 T as well. In this study we investigated the possibility of detecting eye-selective V1 activity using high-resolution GE-EPI fMRI at 3 T and sub-millimeter resolution fMRI at ultrahigh 9.4 T magnetic field strengths with acquisition parameters optimized for each field strength. High-resolution fMRI at 9.4 T also allowed us to examine the eye-selectivity responses across the cortical depth, which are expected to be strongest in the middle layers.

We observed a substantial increase in the percentage of eye-selective voxels, as well as a doubling in run-to-run consistency of eye preference at ultrahigh field compared to 3 T. We also found that across cortical depth, eye selectivity increased towards the superficial layers, and that signal contrast increased while noise remained nearly constant towards the surface. The depth-resolved results are consistent with a distortion of spatial specificity of the GE-EPI signal by ascending venules and large draining veins on the cortical surface. The effects of larger vessels cause increasing signal amplitude, but also displacement of the maximum BOLD signal relative to neural activity.

In summary, our results show that increasing the spatial resolution, reduced partial volume effects, and improved sensitivity at 9.4 T allow for better detection of eye-selective signals related to ocular dominance columns. However, although ultrahigh field yields higher sensitivity to the ocular dominance signal, GE-EPI still suffers from specificity issues, with a prominent signal contribution at shallow depths from larger cortical vessels.

1. Introduction

Recent advances in MRI technologies, in particular the availability of scanners with increasingly higher magnetic fields, allow increasing the spatial resolution of human neuroimaging studies to the submillimeter range. This, in turn, allows noninvasively measuring brain activity generated by mesoscopic functional units within the human cerebral cortex, such as the cortical columns and layers, which, up to now, has been possible primarily in invasive animal studies (Dumoulin et al., 2018).

Human ocular dominance columns (ODCs) are patches of neurons in the primary visual cortex that preferentially respond to the stimulation of one eye (Adams et al., 2007). They represent one of the most popular mesoscopic structures of the cortex that have been examined at ultrahigh...
field. Studies imaging the ocular dominance signal so far focused primarily on mapping out the columnar pattern, either as a proof of principle (Cheng, 2012; Cheng et al., 2001; Menon et al., 1997; Nasr et al., 2016), as a means to benchmark or compare the performance of different acquisition protocols (Feinberg et al., 2016; Yacoub et al., 2007), experimental paradigms (Goodyear and Menon, 2001) or analysis strategies (Kemper et al., 2018), or to study the organization of the ODCs in patients with relevant pathologies, like amblyopia (Goodyear et al., 2002) or achiasma (Olman et al., 2018).

Different fields of neuroscience are not only interested in the mapping of the mesoscopic structures, but also in additionally studying their functional properties or reorganization. Such research questions often rely on the ability to reliably locate those structures as a first step. While some of them may require mapping the exact layout of the columns, merely a reliable detection of voxel biases using univariate or multivariate analyses may be sufficient for others (Chaimow et al., 2017). For example, reliable identification of voxels biased to either the left or the right eye allows examining their responses during a binocular rivalry experiment, providing strong evidence for the involvement of the thalamus and the early visual cortex in binocular rivalry (Haynes et al., 2005). Furthermore, voxel biases toward either eye can be examined before and after monocular deprivation to study the neural plasticity mechanisms of the visual system (Binda et al., 2018).

The majority of previous studies reporting eye-selective activity in V1 used magnetic field strengths of 4 T and higher (Cheng et al., 2001; Feinberg et al., 2016; Kemper et al., 2018; Menon et al., 1997; Nasr et al., 2016; Yacoub et al., 2007). Even at ultrahigh magnetic fields such as the 7 T, a whole scanning session or even averaging over several scanning sessions was required to reliably map the ODC layout (Nasr et al., 2016; Yacoub et al., 2007). At a more conventional magnetic field strength of 3 T, signals related to cortical columns have primarily been reported using multivariate techniques (Chaimow et al., 2011; Haynes and Rees, 2005a, 2005b; Kamitani and Tong, 2005; Shmuel et al., 2010). However, there have been reports of eye-selective activations at a single-voxel level at 3 T as well, not only in V1 but also at the level of LGN (Haynes et al., 2005).

When the acquisition parameters were optimized for this purpose, as little as two runs (5.5 min each) were required to reliably detect eye-selective signals in more than 50 percent of all V1 voxels.

In the present study we aimed to determine the degree to which eye-specific signals can be detected using state-of-the-art and widely available high-resolution (1.5 mm isotropic) imaging at 3 T field strength and to compare this to results obtained using ultrahigh field imaging with 9.4 T field strength that allowed for even higher spatial resolution (0.8 mm isotropic). Our choice of acquisition parameters, and particularly the resolution, was dictated primarily by the success of previous studies discussed above. In addition, we aimed to determine how the fMRI-based ocular dominance signal varied across cortical depth, given that eye-selectivity with sub-millimeter resolution at 9.4 T and that their selectivity is dominated by a strong contribution of the large draining veins on the pial surface and, potentially, ascending intracortical venules. Both types of veins are expected to displace the peak of ODC BOLD signal away from the peak of the underlying neural activity.

2. Methods

2.1. fMRI experiment at 3 T

2.1.1. Participants

Twenty-one healthy adult volunteers (mean age 28.4, range 23–33, 9 females) took part in the 3 T experiment. All had normal or corrected-to-normal vision and no history of neurological impairments. Subjects were screened for MRI-contraindications and hand signed written informed consent. The study was conducted according to the Declaration of Helsinki and was approved by the ethics committee of the University Clinic Tübingen.

2.1.2. Stimulus and experimental procedures

Visual stimulation was generated using MATLAB 2013a (MathWorks, Natick, MA) with the Psychophysics Toolbox 3 extensions (Brainard, 1997; Pelli, 1997) on a Windows computer and presented using a gamma-corrected projector (1920 × 1080 resolution, 60 Hz) viewed at a distance of 80 cm. Independent stimulation of each eye was achieved using a prism stereo setup tailored to the 3 T scanner (Scherger, 2009). In short, participants were wearing goggles with prism lenses in front of each eye. Each prism lens displaced the field of view of the eye inwards by a quarter of the display width thereby leading to the two halves of the screen falling onto the corresponding retinal locations of each eye. To separate the field of view of the two eyes, a wooden plate dividing the length of the scanner bore was placed vertically between the patient table and the screen, and a piece of cardboard was placed between the patient mirror and the wooden plate. The resulting visual field of view was 13° × 16.5° per each eye. Visual stimuli consisted of three experimental conditions, presented in a block-design paradigm: contrast-reversing checkerboard (flickering at 6 Hz) shown to either the left or the right eye, or a gray screen with mean luminance matching that of the checkerboard. The checkerboard conditions alternated with the gray screen condition, starting randomly with either the left or the right eye. Each block lasted 18 s, yielding a total run time of 6 min (5 repetitions of left-/right-eye stimulation blocks and 10 repetitions of the baseline blocks). To ensure fixation participants had to view a stream of letters presented at the center of the screen and to perform a one-back task by pressing a button with an index finger of their right hand. Each run was repeated 4 times, with the exception of one subject, where one run had to be discarded for technical reasons, resulting in 3 runs in total. Only the first three runs were analyzed for all subjects in order to match the overall run number to that of the 9.4 T experiment (see below).

To minimize head movement of the participants, their heads were stabilized with cushions. In addition, participants were informed that in this experiment data would be acquired with particularly high spatial resolution, so even small head movements can compromise the results. They were asked to keep as still as possible during the experiment.

2.1.3. Data acquisition

Experiments were performed on a Siemens Prisma-fit system with a 64-channel head/neck coil array (Siemens, Erlangen, Germany). In 17 out of 21 subjects we used a Siemens syngo ZOOMit EPI sequence with 1.5 mm isotropic resolution (TR/TE/ESP/matrix/GRAPPA = 2000/35/0.7/100 × 60 × 24/R = 2). In the remaining four subjects, BOLD signal was measured using a full field of view 2D GE-EPI using the following parameters: 1.5 mm isotropic resolution, TR/TE/ESP/matrix/GRAPPA = 1720/30/0.93/128 × 128 × 24/R = 3. The resolution was chosen based on the success of a previous study by Haynes et al. (2005).

Structural scans were acquired in each session for cortical surface reconstruction. The parameters followed the ADNI protocol and were identical for all participants (1 mm isotropic resolution, TR/TE/TU/FA/matrix = 2000 ms/3.06 ms/1100 ms/9°/232 × 256 × 192).
2.1.4. Data analysis

2.1.4.1. Processing pipeline. All analyses were performed using FreeSurfer 6.0, FS-FAST (Fischl, 2012) and MATLAB 2015b (Mathworks, Natick, Massachusetts, USA). The reconstruction of gray-white and gray-CSF surfaces for the 3 T data was performed automatically using the FreeSurfer recon-all stream at the default 1 mm isotropic resolution. The correctness of the reconstruction was ensured by visual inspection of the surface alignment to the GM-CSF and GM-WM boarder in the TI-weighted data. In addition to the main reconstruction stream, we also predicted the retinotopic organization (eccentricity and polar angle) of the early visual areas (V1, V2, and V3) in each subject using FreeSurfer and the retinotopic surface template provided by Benson et al. (Benson et al., 2014, 2012). Functional data were motion-correction and co-registered to the anatomical scan using boundary-based registration (Greve and Fischl, 2009) with 6 degrees of freedom. Omitting slice-time correction for our TR and block duration is not expected to substantially affect the magnitude of beta estimates (Sladky et al., 2011), but may help to avoid additional interpolation, which, in combination with subject motion, leads to signal blurring across neighboring voxels (Polimeni et al., 2018). Therefore, we omitted the slice-time correction step. No smoothing was applied to the data in the main analysis. Functional analysis was performed in voxel space. We performed a standard voxel-wise GLM analysis using “left eye”, “right eye” and “baseline” as regressors of interest. Scanner drifts and run-specific offsets were included as nuisance regressors.

2.1.4.2. ROIs and measures of interest for ocular dominance. Regions of interest (ROIs) for V1, V2 and V3 were created by identifying fMRI voxels located within the gray matter of V1, V2 and V3 surface area in each hemisphere. Although the ocular dominance signal is primarily expected in V1, we included V2 and V3 as control regions for comparison to ensure the specificity of our results to V1. The eccentricity estimates from the retinotopic template were used to include only those parts of the visual field that had been stimulated in our experiment (circular region within the central 7.5° of eccentricity). Voxels within these ROIs were subsequently used to calculate the following measures of interest: (1) percentage of eye-selective voxels within each area (defined as the number of eye-selective voxels above the statistical threshold of z = 1.96, corresponding to the p-value of 0.05 uncorrected (two-tailed test), divided by the total number of voxels in an ROI and multiplied by 100) and (2) run-to-run consistency of eye selectivity of voxels within each area (defined as average Fisher-transformed correlation of eye selectivity between all pairs of individual runs). For the correlation analysis, we used signed contrast estimates of the contrast “left eye versus right eye” as regressors of interest. Scanner drifts and run-specific offsets were included as nuisance regressors.

2.1.4.3. Control analyses for potential confounds of within-scanner effects. In order to determine to which extent differences between areas V1–V3 for our two measures of interest can be explained by other factors like differences in temporal signal-to-noise ratio (tSNR), differences in the overall responsiveness of an area to visual stimulation, or differences in the total number of voxels within an area, we also conducted the same statistical analysis for the three control measures: (1) tSNR, estimated as the average beta value of the run-specific offset in the GLM model (an estimate of the mean BOLD signal) divided by the standard deviation of the residuals (an estimate of the temporal noise), (2) BOLD contrast, expressed as percent signal change of the checkerboard response compared to baseline (estimated as the contrast estimates for the contrast “checkerboard versus baseline”, divided by the average beta value for the run-specific offsets in the GLM model and multiplied by 100), and (3) overall number of voxels within each ROI.

To estimate how much of the observed effect in visual ROIs may be attributed to pure noise, we analyzed the data within an additional control ROI that consisted of the white matter voxels. To create the latter ROI, we mapped the structural segmentation results into the functional space and eroded the resulting white matter functional volume by one voxel. The results obtained from the white matter mask were then compared with those obtained in V3 (as an area with smallest effects) using a paired t-test.

To ensure that ocular dominance signal observed in V2 and V3 can’t be explained by partial ROI overlap with V1 (e.g. due to “kissing gyri”), we conducted an additional control analysis in which we excluded those functional voxels that got assigned to more than one ROI (e.g. voxels at the border of the regions or voxels at the opposite banks of the sulcus that belong to different visual areas).

2.2. fMRI experiment at 9.4 T

2.2.1. Participants

Six adult healthy volunteers participated in the 9.4 T experiment (mean age 27.7, range 22–34, 4 females). One of the 9.4 T subjects already participated in the 3 T experiment, while 5 subjects were newly recruited. Subjects were screened for MRI-contraindications and additionally undergone a personal conversation with a medical doctor, after which they signed written informed consent. The study was conducted according to the Declaration of Helsinki and was approved by the ethics committee of the University Clinic Tübingen.

2.2.2. Stimulus and experimental procedures

Visual stimulation for the 9.4 T experiment was similar to that of the 3 T and was presented using a gamma-corrected projector (1920 × 1080 resolution, 60 Hz) viewed at a distance of 142 cm. Independent stimulation of each eye was achieved using a prism stereo setup as described above but tailored to the 9.4 T scanner. The resulting visual field of view was 7.5° × 8.4° for each eye. To separate the fields of view of the two eyes, an opaque black curtain was attached to the top of the scanner bore between the patient table and the screen. The visual stimulation and the paradigm parameters were identical to that of the 3 T experiment (18-s blocks of either the left or the right eye stimulation alternating with 18-s blocks of gray screen). For the 9.4 T experiment each run could only be repeated three times due to time constraints. Similar measures were taken to minimize subject head motion as in the 3 T experiment.

2.2.3. Data acquisition

Experiments were performed on a 9.4 T whole-body MRI scanner (Siemens, Erlangen, Germany) using a custom-built head coil with a 16-element dual row transmit array and a 31-element receive coil array (Shajani et al., 2014). BOLD signal was measured using 2D GE-EPI with the following parameters: 0.8 mm isotropic resolution, TR/TE/ESP/matrix/GRAPPA/partial Fourier = 1810/25/1.01/188 × 188 × 40/R = 5/7/8, FLEET autocalibration (Polimeni et al., 2016). The resolution was chosen based on the success of a previous study by Nasr et al. (2016). Structural scans for cortical surface reconstruction were acquired in a separate session using an MP2RAGE sequence (Marques et al., 2010) adapted for 9.4 T (Hagberg et al., 2017) with the following parameters: 0.6 mm isotropic resolution, inversion TR/readout
TR/TE/TI1/TI2/FA1/FA2/matrix/GRAPPA/partial Fourier = 6s/6 ms/800 ms/2000 ms/5°/9°/352 × 352 × 256/R = 2 × 2/6/8. MP2RAGE data was reconstructed offline (Marques et al., 2010) using custom software developed in MATLAB.

2.2.4. Data analysis

Cortical surface reconstruction from the 9.4 T structural scan was performed at the native 0.6 mm isotropic resolution using the FreeSurfer’s high-resolution pipeline (Zaretskaya et al., 2018). Surfaces were visually inspected to ensure there are no major errors and a few minor errors were corrected (see Supplementary Figure 1 for the surface reconstruction quality). Functional data preprocessing was identical to 3 T, except for the coregistration step. To compensate for potential distortions in the EPI images at ultrahigh field, the boundary-based registration was restricted to only align visual ROIs V1–V3 in each hemisphere, and 9 degrees of freedom were used in the registration instead of 6 (Polimeni et al., 2018). We explicitly decided against distortion correction for the following reasons. First, using a very high in-plane acceleration (GRAPPA R = 5) helped to reduce distortions due to B0 inhomogeneities. Second, our pilot tests did not reveal any major differences in the volume alignment between our approach and the topup method (Andersson et al., 2003) within our ROIs. Third, the unwarping procedure involved in distortion correction introduces blurring in the functional data, which is not desirable in high-resolution experiments (Rause et al., 2020; Polimeni et al., 2018). In our experiment, retaining signal independence in neighboring voxels is particularly critical because neighboring voxels may have opposite ocular preference, and blurring would reduce their detectability. Statistical analysis was identical to that of the 3 T experiment.

2.3. Comparison between 3 T and 9.4 T data

To directly compare the increase in ocular dominance detectability between the 3 T and the 9.4 T data, we conducted a two-way mixed effects ANOVA with one within-subject factor “brain area” (V1, V2, V3) and one between-subject factor “scanner” (3T, 9.4 T). The mixed effects ANOVA was conducted using two variables of interest: percent of eye-selective voxels and run-to-run consistency. A significant area-by-scanner interaction should indicate a 9.4 T advantage in ocular dominance selectivity detection. We also conducted the same ANOVA using the BOLD contrast as a dependent measure (the only control measure that could explain within-scanner effects, see Results for details). To make sure that the difference in sample sizes between scanners does not affect the statistical results, we also conducted Levene’s test of equality of error variance for all ANOVAs containing between-subject factors (Mycroft et al., 2002), none of which showed a significant effect at p < 0.05. In order to test which differences drive the interaction we performed post-hoc t-tests for unpaired samples between the scanners separately for each area.

In order to determine to which extent high resolution is indeed important for ODC detection, we conducted an additional analysis of the 9.4 T data. We progressively smoothed the preprocessed fMRI data with 3D Gaussian kernels of different FWHMs ranging from 1 to 2.5 mm in steps of 0.25 mm, each time recomputing the GLM. This resulted in eight 9.4 T datasets, one original and seven smoothed. We then examined how our measures of interest (% of eye-selective voxels and run-to-run consistency) changed with smoothing-induced loss of spatial resolution. In addition, we also examined the change in 1) ODC contrast (absolute difference of beta estimates between the left and the right eye, averaged across voxels), 2) temporal noise (standard deviation of the GLM residuals, averaged across voxels) and 3) ODC contrast-to-noise ratio (first measure divided by the second measure, averaged across voxels), which ultimately determines the outcome of the two measures of interest. To statistically test the smoothing effect, we compared the values derived from the unsmoothed data with a smoothing level corresponding to a Gaussian with 1.25 mm FWHM, which should be roughly equivalent to our 3 T voxel size (assuming voxel FWHM and kernel FWHM add in quadrature, 0.8 mm voxel smoothed with a 1.25 mm FWHM kernel results in an effective voxel size of 1.48 mm).

2.3.1. Control analyses for potential confounds of between-scanner effects

To exclude the possibility that the differences between the 3 T and the 9.4 T results can be explained by different amount of subject movement, we characterized each subject’s movement throughout the experiment from the realignment parameters as the median (across all analyzed images) vector motion of the center of rotation of each image relative to the target image. To determine whether the amount of head motion is an important factor affecting the detection of eye-selective activity, we additionally correlated the amount of head motion with the percentage of voxels above the threshold and with the individual run-to-run consistency across subjects. In order to determine whether differences in ocular dominance signal between the scanners is due to differences in the amount of movement, we compared the amount of movement between the two participant groups using a t-test for unpaired samples.

In addition, to make sure the differences in the extent of visual stimulation between the scanners (13° at 3 T versus 7.5° at 9.4 T) do not affect our result, we repeated the scanner comparison using control ROIs that included only central 6° of the visual field for both scanners. This equated the sub-regions within V1–V3 that were analyzed across experiments.

Finally, to rule out the possibility that the differences in the % of eye-selective voxels between scanners can be explained by the differences in the rate of false positives, we repeated the between-scanner comparison for the % of eye-selective voxels computed using an FDR-adjusted threshold corresponding to p < 0.05 corrected.

2.4. Depth-dependent analysis of 9.4 T data

For the cortical depth analysis, the two surfaces of the gray matter sheet created during the FreeSurfer surface reconstruction were used to generate nine additional intermediate surfaces spanning the gray matter thickness and spaced at equal distances, yielding 11 depth values ranging from 1 (deepest) to 0 (most superficial) (Polimeni et al., 2010). Each of these 11 cortical surfaces was projected into the functional volume space. Voxels intersecting surface locations occupied by each visual area were labeled with the corresponding depth value ranging from 0 (deepest voxels) to 1 (most superficial voxels), yielding a set of 11 depth ROIs for each visual area, separately for each hemisphere.

The known anatomy and electrophysiology of the primary visual cortex suggests that eye-selective responses should be highest in the middle layers of V1 (Hubel and Wiesel, 1968). To test this hypothesis, we examined three different measures of ocular dominance selectivity across the cortical depth. In addition to computing the percent significant voxels and the run-to-run consistency as was performed for our main analysis, we also determined 1) ODC contrast (absolute difference of beta estimates between the left and the right eye, averaged across voxels), 2) the noise (standard deviation of the residuals from the GLM model, averaged across voxels) and 3) the ODC contrast-to-noise ratio (the first measure divided by the second measure, averaged across voxels). This was performed separately for each cortical depth. Since we observed a gradual increase of the first three measures towards the cortical surface (see Results section for details), we quantified this increase by fitting a line to each subject’s depth profile and statistically comparing the slopes against zero using a two-tailed one-sample t-test.

To test our hypothesis that eye selectivity in the superficial voxels stems from larger vessels, we conducted the following additional analysis. We sampled the spatial distribution of eye selectivities (unthresholded t-statistic maps for the contrast “left eye versus right eye”) at different cortical depths by projecting the 3D map onto each of the subsurfaces used in our depth analysis. We then smoothed these statistical maps on the surface by convolving each of them with a Gaussian with increasing FWHM (0–4 mm in steps of 1 mm). This resulted in a 5 × 11
matrix of 5 smoothing levels and 11 depths. We then recomputed our measure of interest (% of eye-selective voxels) for each cell of the matrix. If the ocular dominance responses closer to the cortical surface stem from larger vessels, eye-selectivity should vary at a coarser spatial scale than the ODCs themselves, as has been shown by Shmuel et al. (2010). Consequently, large smoothing kernels should destroy the fine-scale ODC pattern of around 1 mm in deeper layers, but have less effect on the results for the superficial layers.

2.5. Data and code availability statement

GLM results, ROIs and code are available from the corresponding author upon request. Raw data may be made available upon request after confirmation from the ethics committee, which is pending.

3. Results

3.1. Individual example

Results for the left hemisphere of an individual subject who took part in both the 3 T (full field of view sequence) and the 9.4 T experiment are shown in Fig. 1. Although a large percentage of eye-selective voxels (9.6%) could be identified in this individual in the 3 T experiment (Fig. 1A), the run-to-run correlation of eye preference across V1 voxels was very low (1st run versus 3rd run Pearson’s \( r = 0.07 \)), suggesting that many voxels in V1 were false positives. In contrast, a larger percentage of voxels (19%) showed statistically-significant eye preference in the 9.4 T experiment (Fig. 1B), and this pattern was correspondingly more consistent across runs (1st run versus 3rd run Pearson’s \( r = 0.43 \)). Please note that the above percentages refer to all voxels included in an ROI irrespective of their responsiveness to visual stimulation.

3.2. Results of the 3 T experiment

The pattern seen in this representative subject is reflected in the group results as well (Fig. 2). The largest percentage of voxels with significant eye preference was found in area V1, followed by V2 and V3 (Fig. 2A and Table 1, first column). This gradual decrease in the percentage of eye-selective voxels was significant (F(1,2,23.9) = 36.07, \( p < 0.001 \), \( \eta^2_p = 0.64 \), all post-hoc t-tests are significant at \( p < 0.001 \) corrected for multiple comparisons using the Bonferroni method). Similarly, area V1 showed the highest run-to-run consistency of voxel preference, followed by V2 and V3 (Fig. 2B and Table 1, second column). Again, this decrease from V1 to V3 was significant (F(1,4,29.6) = 44.51, \( p < 0.001 \), \( \eta^2_p = 0.69 \), all post-hoc t-tests are significant at \( p \leq 0.02 \) Bonferroni-corrected).

The above results could be explained neither by the differences in tSNR across regions (F(5) = 3.19, \( p = 0.06 \), with V2 having the highest tSNR, followed by V3 and V1), nor by the number of voxels included in each ROI (F(2, 40) = 27.80, \( p < 0.001 \), \( \eta^2_p = 0.58 \), with V3 being the largest ROI, followed by V1 and V2). Note that although the ROIs differed significantly in the number of voxels, the pattern of differences did not correspond to the pattern of the main results (Table 2, first and third columns). However, we observed a significant increase in the BOLD contrast across areas (F(1.2, 23.1) = 81.98, \( p < 0.001 \), \( \eta^2_p = 0.80 \) that could potentially contribute to the observed main results: V1 showed the strongest response, followed by V2 and V3 (all post-hoc tests significant at \( p < 0.001 \) Bonferroni-corrected). Hence, although our results suggest that eye-selectivity signal may be detectable at 3 T, we cannot rule out that the observed ocular dominance responses are attributable to the overall differences in responsiveness to visual stimulation.

We further examined to what extent the eye-selectivity effects in areas V2 and V3 can be explained by pure noise and false positives. We compared the two measures of interest obtained in V3 to those derived from the white matter mask (WM control). The percentage of eye-selective voxels was significantly higher in V3 compared to the WM control ROI (t(20) = 3.58, \( p = 0.002 \), \( d = 0.78 \), as was the run-to-run consistency (t(20) = 2.87, \( p = 0.009 \), \( d = 0.63 \)), so the effect in V3 (and even more so in V2) is unlikely to be due to noise alone. In addition, we could rule out that eye-selectivity effects in areas V2 and V3 are due to signals generated in V1. After excluding voxels that got assigned to more than one area, there were no major changes in the observed values (Table 1, round brackets). On average, such voxels comprised only 0.33 ± 0.023% of all voxels of the three ROIs (mean ± SE across 3 T subjects).

3.3. Results of the 9.4 T experiment

Although the voxel volume in the 9.4 T experiment was 6.6 times smaller (0.51 mm³ versus 3.38 mm³), qualitatively the results were similar to those at 3 T. The largest percentage of eye-selective voxels was found in V1, followed by V2 and V3 (Fig. 2A and Table 1, first column). The difference in the percentage of eye-selective voxels across areas was significant (F(2, 10) = 16.29, \( p = 0.001 \), \( \eta^2_p = 0.77 \)). Area V1 differed significantly from V2 and V3 (t(5) = 4.64, \( p = 0.017 \) Bonferroni-corrected), and from V3 (t(5) = 4.12, \( p = 0.027 \) Bonferroni-corrected), but there was no difference between V2 and V3 (t(5) = 2.57, \( p = 0.05 \) uncorrected). Similarly, area V1 showed the highest run-to-run consistency of voxel selectivity, followed by V2 and V3 (Fig. 2B and Table 1, second column). The difference in the run-to-run consistency across areas was also significant (F(2, 10) = 11.88, \( p = 0.002 \), \( \eta^2_p = 0.70 \)). The comparison of V1 with V2 did not reach statistical significance (t(5) = 3.19, \( p = 0.073 \) Bonferroni-corrected), but the comparison with V3 did (t(5) = 3.65, \( p = 0.044 \) Bonferroni-corrected). V2 and V3 did not differ significantly from each other (t(5) = 3.05, \( p = 0.085 \) Bonferroni-corrected).

Fig. 1. Comparison of 3 T and 9.4 T results in an individual participant. Voxel-wise GLM results (z-values) comparing brain activity during left eye stimulation and right eye stimulation are projected on an inflated cortical surface of the left hemisphere. White outline represents the predicted location of the stimulated part of the primary visual cortex (V1). Insets show scatter plots of voxel selectivity and a corresponding Pearson’s correlation coefficients in eye preference between the 1st and the 3rd runs prior to Fisher-transform. Voxels are sampled at 50% of the cortical depth and thresholded at \( p < 0.05 \) uncorrected.
Also here, all overlapping voxels comprised only a small fraction (0.09 ± 0.004%, mean ± SE across 9.4 T subjects) and their exclusion did not substantially change the observed values (Table 1, round brackets).

As for the 3 T analysis, we examined to what extent the eye-selectivity effects in areas V2 and V3 can be explained by pure noise and false positives. Also in the 9.4 T experiment the percentage of eye-selective voxels was significantly higher in V3 compared to WM control ROI (t(5) = 5.55, p = 0.003, d = 2.26), as was the run-to-run consistency (t(5) = 2.87, p = 0.035, d = 1.17). Hence, eye-selectivity in V3 is unlikely to be due to noise alone. Also here, all overlapping voxels comprised only a small fraction (0.09 ± 0.004%, mean ± SE across 9.4 T subjects) and their exclusion did not substantially change the observed values (Table 1, round brackets).

### Table 1
Average and SEM values for each of the measures of interest in a 3 T and a 9.4 T experiments. Numbers in round brackets indicate the same values, but after excluding voxels assigned to more than one area.

|                  | 3 T                     | 9.4 T                  |
|------------------|-------------------------|------------------------|
| % eye-selective voxels | V1: 13.7 (13.1)         | V1: 15.3 (15.2)        |
|                  | V2: 10.7 (10.7)         | V2: 11.0 (11.0)        |
|                  | V3: 8.7 (8.9)           | V3: 10.7 (10.7)        |
|                  | WM: 5.4                 | WM: 5.0                |
| Run-to-run consistency | V1: 0.1 (0.1)          | V1: 0.05 (0.05)        |
|                  | V2: 0.1 (0.1)           | V2: 0.05 (0.05)        |
|                  | V3: 0.1 (0.1)           | V3: 0.05 (0.05)        |
|                  | WM: 0.05                | WM: 0.05               |

3.4. Main results

The above results suggest that for both experiments, the highest and most reliable ocular dominance signal is found in V1, but also that ocular dominance selectivity is found at more cortical locations, and with higher reliability at 9.4 T compared to 3 T. Indeed, the two-way ANOVA with scanner as a between-subject factor showed a significant effect of “area” for both measures of interest (percent of eye-selective voxels F(1, 12, 30.0) = 59.43, p < 0.001, η² = 0.70, observed power = 1; run-to-run consistency: F(1, 33.7) = 54.63, p < 0.001, η² = 0.69, observed power = 1). The main effect of “scanner” was significant for the run-to-run consistency (F(1, 25) = 9.44, p = 0.005, η² = 0.27, observed power = 0.84), but did not reach significance for the percent of eye-selective voxels (F(1, 25) = 4.10, p = 0.05, η² = 0.14, observed power = 0.49). Importantly, there was a significant area-by-scanner interaction for both measures (percent of eye-selective voxels: F(1, 2, 30.0) = 9.42, p = 0.003, η² = 0.27, observed power = 0.89; run-to-run consistency: F(1, 3, 33.7) = 7.64, p = 0.005, η² = 0.23, observed power = 0.84). Post-hoc t-tests showed significantly higher values for 9.4 T compared to 3 T for both measures of interest in V1 (percent of eye-selective voxels p = 0.005 Bonferroni-corrected; run-to-run consistency p = 0.003 Bonferroni-corrected), but not for V3 (both p ≥ 0.18 Bonferroni-corrected), with mixed results for V2 (percent of eye-selective voxels p = 0.15 Bonferroni-corrected; run-to-run consistency significant: p = 0.004 Bonferroni-corrected).

3.4.2. Control analyses

Although differences in the overall checkerboard response between areas can explain the main effect of area, they cannot explain the above interaction effect. ANOVA using the BOLD contrast as a dependent variable showed a significant main effect of area (F(1, 28.4) = 73.77, p < 0.001, η² = 0.74, observed power = 1) and a significant main effect of scanner (F(1, 25) = 22.88, p < 0.001, η² = 0.48, observed power = 1). Crucially, there was no area-by-scanner interaction (F(1, 28.4) = 2.39, p = 0.12).
observed power ($p = 0.13$), indicating that higher field strength boosted the BOLD contrast in V1 and V3 by the same amount. Post-hoc tests revealed that the checkerboard response was significantly larger at 9.4 T compared to 3 T in V1 ($p = 0.004$ Bonferroni-corrected) and in V2 ($p < 0.001$ Bonferroni-corrected), but also in V3 ($p < 0.001$ Bonferroni-corrected).

The difference between 3 T and 9.4 T could not be explained by the differences in the amount of head movement, as in both experiments median vector motion was slightly above 1 mm (mean and S.E.M for 3 T: $1.30 \pm 0.16$ mm; 9.4 T: $1.36 \pm 0.26$ mm) and did not differ statistically ($t(25) = -0.2, p = 0.84$). Surprisingly, the amount of head motion in our 3 T subject sample, which was large enough to compute correlation, affected neither the ability to detect eye-selective voxels (correlation of head motion with percentage of eye-selective voxels: $R = 0.17, p = 0.76$), nor the run-to-run consistency (correlation of head motion with run-to-run consistency: $R = 0.23, p = 0.33$).

We could rule out that the significant area-by-scanner interaction observed for our measures of interest can be explained by the differences in the extent of visual stimulation between scanners. ANOVAs with measures of interest derived from the control analysis with ROIs in which the extent of V1-V3 was identical in both scanners (central 6° of eccentricity) yielded comparable effects. Specifically, there was a significant main effect of area ($F(1.2, 30.0) = 59.43, p < 0.001, \eta^2_p = 0.70$, observed power = 1) and a significant area-by-scanner interaction ($F(1.2, 30.0) = 9.42, p = 0.003, \eta^2_p = 0.27$, observed power = 0.88) for the percentage of eye selective voxels, while the effect of scanner did not reach significance ($F(1, 25) = 4.10, p = 0.05$). There was a significant main effect of area ($F(1.4, 33.7) = 54.63, p < 0.001, \eta^2_p = 0.69$, observed power = 1), significant main effect of scanner ($F(1, 25) = 9.44, p = 0.005, \eta^2_p = 0.27$, observed power = 0.84) and a significant area-by-scanner interaction ($F(1.4, 33.7) = 764, p = 0.005, \eta^2_p = 0.23$, observed power = 0.84) for the run-to-run consistency.

To rule out that the effects related to the percentage of eye-selective voxels were due to the differences in false positive rate between the scanners, we repeated the analysis after correcting the p-values for multiple comparisons using the FDR method ($p < 0.05$) within each ROI. Percentages of eye-selective voxels after the FDR correction are shown in Supplementary Figure 2. Although the percentages were reduced overall, the main pattern of results remained the same. There was a significant main effect of area ($F(1.2, 28.7) = 33.84, p < 0.001, \eta^2_p = 0.58$, observed power = 1), a significant area-by-scanner interaction ($F(1.2, 28.7) = 12.05, \eta^2_p = 0.33$, observed power = 0.94), while the main effect of scanner did not reach significance ($F(1, 25) = 2.39, p = 0.14$). Again, 9.4 T was different from 3T in V1 ($p = 0.01$ Bonferroni-corrected), but not in V2 and V3 (both $p > 0.3$ Bonferroni-corrected).

To determine the role of high resolution in the observed 9.4 T advantage, we conducted an additional smoothing analysis of 9.4 T data. The outcome of smoothing the 9.4 T data is shown in Fig. 3. We observed that both the percentage of eye-selective voxels and the run-to-run consistency increased with larger smoothing kernels (Fig. 3A and B). The ODC contrast progressively decreased with smoothing (Fig. 3C), which suggests that the ODC signal we measure is present at a fine spatial scale, and the eye selectivity is lost because smoothing leads to averaging signals from neighboring columns with opposite selectivity. At the same time, smoothing reduces noise (Fig. 3D). Since noise reduction is stronger

![Fig. 3](image-url) Effect of smoothing on 9.4 T data in V1. (A) Percentage of voxels showing significant eye preference in V1 and (B) Run-to-run consistency of voxel selectivity in V1 for smoothing with Gaussian kernels of different FWHM. (C) ODC contrast, defined as the absolute difference of beta estimates for the left and right eye, averaged across all voxels within the V1 ROI. (D) Noise, defined as the standard deviation of the GLM residuals, averaged across voxels within the V1 ROI and (E) ODC contrast-to-noise ratio, defined as contrast divided by noise, averaged across voxels within the V1 ROI. Gray arrows indicate the smoothing level approximately equivalent to the 3 T voxel size (1.25 mm FWHM). *$p < 0.05$; **$p < 0.05$; ***$p < 0.001$ for a two-tailed paired $t$-test between no smoothing and 1.25 mm FWHM. Error bars indicate the standard error of the mean across six subjects.
than the resolution loss, the outcome is an increase in ODC-CNR (Fig. 3E),
which ultimately determines the outcome for our measures of interest.
Given that in our 9.4 T and 3 T experiments the tSNR was comparable
(and even slightly higher at 3 T), we suggest that the main advantage of
the 9.4 T in detecting the ODC signal is due to higher spatial resolution
(at comparable SNR).

3.5. Depth-dependency of the ocular dominance signal

Although V1 anatomy and physiology predict that the ocular domi-
nance response should be strongest in the middle layers of V1 due to the
input from the monocular LGN layers, we found a gradual increase in
both measures of interest towards the superficial layers of V1 (Fig. 4).
The statistical comparison of the slopes of this gradual increase showed
that it is significantly different from zero for both measures in each area
(Table 3). This gradual increase in gradient-echo BOLD activation across
depths is consistent with the well-known bias of pial veins and ascending
venules (Bause et al., 2020; De Martino et al., 2013; Koopmans et al.,
2010; Moerel et al., 2018; Polimeni et al., 2010). In contrast to the volume
smoothing analysis of the 9.4 T data (Fig. 3), which showed that gradual noise reduction can lead to increase in detectability of
eye-selective voxels, the depth-dependent effects were driven primarily
by the increase in the contrast, with little or no change in the noise levels.

To determine whether our ODC signals exist at a fine spatial scale of
the true ocular dominance columns or at a coarse spatial scale of larger
vessels, we smoothed the pattern along the cortical surface at different
depths. The results of this analysis are shown in Fig. 5. We observed that
at deep locations all smoothing kernels reduced the percentage of eye-
selective voxels to the level found in the white matter, but that they
had much less effect closer to the surface (Fig. 5B). Even smoothing
kernels as large as 3 mm FWHM are well above the white matter level in
the superficial layers, suggesting a contribution from the coarse spatial
scale of blood vessels.

4. Discussion

Here we investigated the influence of MR field strength and associ-
ated distinct spatial resolutions on isolating eye-specific signals in the
early visual cortex. We found that the increase in spatial resolution and
functional CNR at 9.4 T allows for better and considerably more reliable
detection of eye-selective signals related to ocular dominance columns.
4.1. Detecting ODC signal at 3 Tesla

Using a large sample of 21 naïve subjects we could confirm that eye selectivity can be detected on a single-voxel level at a conventional field strength of 3 T. On average, 15% (range 6–28%) of V1 voxels showed a significant eye preference after 18 min of scanning time, and there is a relatively small but significant run-to-run consistency of the eye selectivity signal, suggesting that the above percentage is not driven by the false positives alone.

Our results appear to be more modest compared to what has been seen...
previously reported for V1 (Haynes et al., 2005). In their study, Haynes et al. have reported that on average 60% of V1 voxels (range 52–71% over four subjects, Supplementary Figure S1 in Haynes et al.) showed eye-selectivity after only 11 min of scanning. What could be the reasons for this discrepancy? One possibility is the difference in hardware or acquisition parameters. In four of our subjects we have tried to match our protocol parameters to those of Haynes et al., but the results of those four subjects were not substantially different from the group. Another possibility is the difference in experimental paradigm. While in the Haynes et al. study one eye was stimulated throughout one run, in our experiment the stimulation of each eye alternated within each run. We deliberately chose the latter option in order to avoid the potential monocular deprivation effects and to reduce adaptation effects. Recent studies show that short-term monocular deprivation of 2 h can significantly affect eye preference of voxels towards the deprived eye (Binda et al., 2018). If such plasticity effects exist on a shorter time scale as well, it is important to determine whether they could enhance or bias eye-selectivity detection. One last possibility is the discrepancy in the number of participants (21 versus 4) – if Haynes et al. used participants that had been selected on the basis of prior fMRI experiments or even of ocular dominance pilot experiments this could account for higher-than-average results. However, if one considers the inter-individual variability in the percentage of eye-selective voxels in our sample, this possibility seems unlikely. Even the value of our best 3T subject with 28% of eye-selective voxels is far lower than the worst subject in the Haynes et al. (2005) study. Hence, if a subject pre-selection contributed to the result discrepancy, then only in combination with other factors discussed above.

4.2. Ultrahigh field advantages and limitations of the comparison

Not surprisingly, the comparison of 3 T and 9.4 T results shows a clear advantage of high spatial resolution in the 9.4 T experiment. One major difference between the two experiments that could potentially bias the comparison is the difference in visual stimulation. The stimulus within the 9.4 T scanner bore was farther away from the subject, which resulted in a smaller visual field of view. As a consequence, our analysis of 3T data included more eccentric regions of the visual areas compared to the 9.4 T data. However, this would not be expected to affect the results: the territory of each eye’s columns is known to be approximately equal within the central 15 degrees of eccentricity (Adams et al., 2007), and in both experiments the stimulation was well within the central 15°. We nevertheless made sure that this difference could not explain our results by conducting an additional control analysis within the central 6 degrees of eccentricity for both scanners. This analysis showed results that are comparable to the main analysis, thus making this bias extremely unlikely.

On the other hand, in our 9.4 T experiment we utilized several newest developments in sequence and software for ultrahigh-field MRI, which were not feasible/accessible in our 3 T experiment. For example, we used the FLEET autocalibration scan for in-plane normalization, which helps to reduce artifacts related to subject motion and respiration during the GRAPPA calibration scan and the run-to-run variability in tSNR (Blazejewska et al., 2017; Polimeni et al., 2016), which, in turn, increases the sensitivity to eye-selectivity. In addition, our high resolution T1-weighted structural scans (0.6 mm isotropic vs. 1 mm isotropic) have likely led to a more accurate cortical surface reconstruction (Zaretskaya et al., 2018), and, as a consequence, more accurate detection of voxels within the V1 gray matter. While the latter factor could have influenced the percent voxels with ocular dominance, it would not have influenced the run-to-run consistency measure. Nevertheless, it is important to point out that there are several factors beyond mere field strength and voxel size that inherently differ between 3 T and 9.4 T data acquisition. The aim here was hence not to attribute the observed difference of ocular dominance measure to field strength only, but to provide a realistic comparison between 3 T and 9.4 T when using optimal methods for each of the approaches.

In a recent modeling work that took into account the properties of the columnar pattern, the point spread function of the BOLD signal, and the imaging parameters, Chaimov et al. predicted the optimal voxel size for detecting columnar structures at 3 T and 7 T (Chaimov et al., 2017). Although our choice of voxel size at 3 T (1.5 mm) was mainly determined by the success of a previous study (Haynes et al., 2005), it roughly matches the predicted optimal voxel size for multivariate detection at 3 T. Assuming the cycle width of the pattern being 1.8 mm (Adams et al., 2007), the optimal voxel size according to Chaimov et al. is between 1.3 and 1.5 mm. For 7 T the predicted optimal voxel size should be between 0.7 and 1 mm. Given that 9.4 T should yield higher SNR (Pohmann et al., 2016), voxel size in our experiment could have been smaller in theory but was not feasible due to technical limitations of gradient performance.

It is remarkable that the tSNR in this study was only 1.2 times lower at the 9.4 T, although the voxel volume was 6.6 times smaller (0.51 mm³ versus 3.38 mm³). The fact that 9.4 T shows an advantage in the detectability of the ODC-related signals despite comparable tSNR suggests that the spatial resolution (and potentially also the predominately extravascular signal origin) is the most important factor for detecting the ODC-related signals. Although it has been shown that larger vessels and other large structures may have a significant eye-bias, which can be detected at a coarser spatial resolution at 3 T (Shmuel et al., 2010), decreasing the resolution to the width of a single ocular dominance column and smaller should increase the probability of any voxel containing predominantly a column of one of the eyes, thereby increasing the percentage of eye-selective voxels overall.

Finally, it is important to point out that our 9.4 T sample size of six participants is quite modest compared to the sample of the 3 T experiment. We therefore took extra precautions while conducting statistical tests to make sure the small sample size for the 9.4 T experiment does not bias the comparison across scanners (by e.g. testing for the equality of error variance). We also ensured that our readers are informed about potential issues with the reliability and replicability of our findings by reporting the effect sizes and observed power of our main tests for both within- and between-scanner effects. Although we generally observed large effects and high power, we would like to emphasize that with increasing availability of the ultrahigh field scanners it is important to use larger sample sizes in order to avoid the reproducibility issues.

Despite the state-of-the-art technological advances, superior SNR and, as a consequence, higher resolution, our ability to detect ocular dominance signal and its consistency on a run-to-run basis still appears to be modest. This fact should be taken into account when planning basic neuroscience and patient studies that require detection of signals related to columnar structures.

4.3. Depth-dependent effects

Our main findings of ocular dominance signal changes across the cortical depth show that eye-selectivity increases towards the superficial cortical locations. We confirmed that this effect is driven primarily by signal amplitude increase towards the cortical surface, which was stronger than the increase in the amount of noise. These findings are consistent with many previous studies showing that the signal in high-resolution GE-EPI is biased towards the larger pial surface veins and intracortical ascending venules (De Martino et al., 2013; Moerel et al., 2018; Polimeni et al., 2010), which is a well-known problem of gradient-echo EPI acquisitions.

Recently, two models have been proposed that allow accounting for the effect of draining veins, resulting in less confounded and more spatially specific laminar profiles (Heinzle et al., 2016; Markweertka et al., 2016). Although such models are very promising, and attempts have been made to apply them to neuroimaging experiments across cortical depth (Marquardt et al., 2018), their application to any depth-resolved analysis is not straightforward. First of all, the models lack validation against ground truth human data. Second, such models account only for the ascending venule effect, but not the effects of the pial
vasculature, which may be the dominating source of bias (Bause et al., 2020). Fourth, even if the first two major biases are eliminated, there are still differences in laminar capillary density across the cortical depth, which may be the source of a secondary bias of depth-dependent fMRI signals (Duvernoit et al., 1981; Koopmans and Yacoub, 2019; Norris and Polimeni, 2019). Finally, several previous studies have shown that differential responses across cortical depth may be detectable even without modeling procedures (Huber et al., 2017; Kok et al., 2016; Marquardt et al., 2018).

Another suggestion for reducing large vessel bias in GE-EPI has been the use of differential experimental paradigms, in which the differential response between two conditions helps to reduce any unspecific activity in the BOLD signal under the assumption that the signal in large draining veins is comparable during the different conditions (Markweierka et al., 2016; Sanchez Panchuelo et al., 2015). Our depth-dependent results show that even for such closely matched stimulus conditions as the stimulation of one versus the other eye that are subtracted to determine eye selectivity, the detected selectivity still contains significant vessel biases originating from locations different from those of the underlying neural signal.

Our depth-dependent results may appear contradictory. Eye selectivity stemming from the ocular dominance columns requires a high degree of spatial specificity, so why do we observe an increase in the ODC contrast closer to the cortical surface, which is dominated by large vessels that should exhibit low spatial specificity? One of the factors that can contribute to such an effect is the previously reported eye-selectivity found in larger vessels (Shmuel et al., 2010). In their study, Shmuel et al. identified the likelihood of each voxel in their high-resolution 7T GE dataset to stem from a vessel and found that voxels with vessel contamination are as predictive of the stimulated eye as the gray matter. Due to the acquisition protocol (anisotropic spatial resolution with a slice thickness spanning the whole of the cortical thickness), their analysis could not resolve responses across depths. Our use of isotropic voxels allowed us to examine the signals across the cortical depth as well, yet we came to a similar conclusion. In our data, eye selectivity is highest at the upper cortical layers where larger vessels are located, and not in the middle layers, the site of the strongest neural eye selectivity in V1. The fact that eye selectivity at the cortical surface in our study is distributed at a coarser spatial scale than what can be expected for realistic ODCs is confirmed by our surface smoothing analysis. Hence, despite the predominantly extravascular signal origin, strong domination of the signals from around the larger vessels may have prevented us from detecting the true neural depth-dependent effects (Uludag et al., 2009).

It is also possible that, although layer 4 of V1 receives monocular LGN inputs, its functional responses do not show increased monocular selectivity compared to other layers, which could be an alternative explanation for our failure to see an increase of eye selectivity in the middle layers. Early deoxy-glucose experiments showed that during monocular stimulation strongest responses are observed in the lower layers 4 and 6, whereas the responses in the superficial layers are weakest (Tootell et al., 1988). Interestingly, although the monocular response was strongest in layer 4, the same layer showed an increased response within the columns of the unstimulated eye, suggesting binocular interactions in this layer. Such an effect could have led to the overall smaller difference between the left and the right eye in the middle layers in our experiment. Single-cell physiology results show the strongest ocular dominance selectivity in layers 4 and 5, but also the weakest responses in the superficial layers (Hubel and Wiesel, 1968). Hence, even if the neural oculart selectivity is not an exclusive property of the middle layers, the eye selectivity signal across cortical depth in our experiment should either decrease towards the superficial layers or at least remain constant, whereas we observe a gradual increase towards the cortical surface, consistent with the venous bias effect.

Another interesting question is why, if the eye selectivity signal is dominated by larger vessels, should the spatial resolution play such an important role, as our 3 T versus 9.4 T comparison shows. One of the reasons may be that a subset of voxels (those located deeper within the gray matter) may exhibit more spatially specific signals (Polimeni et al., 2010). Another reason is that submillimeter resolution at 9.4 T helps detecting extravascular signals from smaller vessels compared to 3 T, and smaller vessels, which drain blood from comparatively smaller cortical regions, may be more likely to exhibit a signal bias towards one of the eyes.

4.4. Ocular dominance selectivity beyond V1: false positives or true results

An interesting effect observed in our study is the presence of eyeselective voxels in areas V2 and V3, both at 3 T and at 9.4 T. The percentage of eye-selective voxels that we report in Fig. 2 is likely to also include false positives. Control analysis within the white matter confirms this, but also suggests that these effects cannot be explained by false positives alone. Moreover, false positives cannot explain the above-zero run-to-run consistency in eye-selectivity. While some of the eye-selectivity in V2 can be explained by small errors in the automatic prediction of the V1–V2 boundary, this is an unlikely explanation for V3, which does not have a border with V1. These results are compatible with primate electrophysiology, which has shown that neurons at different processing stages beyond V1 display biases in their ocular preferences (Burkhalter and Van Essen, 1986; Maunsell and Van Essen, 1983; Uka et al., 2000). The present data suggest that clusters of neurons with monocular biases beyond V1 may be sufficiently large to also bias BOLD responses of voxels, or that nonlinearities in hemodynamic coupling allow also small neuronal clusters to influence a voxel-level response. More studies are needed for a better understanding of these effects.

5. Conclusion

Although both the high-resolution acquisition at 3 T and the submillimeter resolution at ultrahigh field 9.4 T allow measuring eyespecific biases related to the ocular dominance columns, sub-millimeter resolution at 9.4 T allows a more reliable detection at more visual cortical locations. Nevertheless, our data suggest that even with unprecedented SNR and high resolution, the detection of eye-selective activity on a single run basis is modest and is not sufficient for mapping of the columnar layout. Furthermore, our data confirm previous results showing that GE-EPI suffers from spatial specificity issues that make it harder to resolve signals related to the layer-specific or column-specific neural activity.

CRediT authorship contribution statement

Natalia Zaretskaya: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Funding acquisition. Jonas Bause: Methodology, Software, Writing - review & editing. Jonathan R. Polimeni: Methodology, Writing - review & editing. Pablo R. Grassi: Software, Investigation. Klaus Scheffler: Resources, Supervision, Writing - review & editing. Andreas Bartels: Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Funding acquisition.

Acknowledgments

The authors would like to thank Johannes Boldt for his help with building the prism stereo setups and Christian Secker, Fatemeh Molaei Vaneghi and Susanna Schiavi for their assistance in data collection. This work was funded by the Centre for Integrative Neuroscience, Tübingen, Germany, German Excellence Initiative grant number EXC307, by DFG Grant BA4914/1-1, and by the Max-Planck Society, Germany. The authors also acknowledge the financial support by the University of Graz, Austria.
Sladky, R., Friston, K.J., Trostl, J., Cannington, R., Moser, E., Windischberger, C., 2011. Slice-timing effects and their correction in functional MRI. Neuroimage 58, 588–594. https://doi.org/10.1016/j.neuroimage.2011.06.078.

Tootell, R.B., Hamilton, S.L., Silverman, M.S., Switkes, E., 1988. Functional anatomy of macaque striate cortex. I. Ocular dominance, binocular interactions, and baseline conditions. J. Neurosci. 8, 1500–1530.

Uka, T., Tanaka, H., Yoshiyama, K., Kato, M., Fujita, I., 2000. Disparity selectivity of neurons in monkey inferior temporal cortex. J. Neurophysiol. 84, 120–132. https://doi.org/10.1152/jn.2000.84.1.120.

Uludağ, K., Müller-Bierl, B., Ugurbil, K., 2009. An integrative model for neuronal activity-induced signal changes for gradient and spin echo functional imaging. Neuroimage 48, 150–165. https://doi.org/10.1016/j.neuroimage.2009.05.051.

Yacoub, E., Shmuel, A., Logothetis, N., Ugurbil, K., 2007. Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. Neuroimage 37, 1161–1177. https://doi.org/10.1016/j.neuroimage.2007.05.020.

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. https://doi.org/10.1016/j.neuroimage.2017.09.060.