

# Current Biology

## Transformation from a Retinal to a Cyclopean Representation in Human Visual Cortex

### Highlights

- Binocular stimuli reveal a change in topographic organization of early visual cortex
- Responses in V1 are best predicted by stimulus position in the two retinal images
- Starting in V2, responses are better predicted by the position in the cyclopean image

### Authors

Martijn Barendregt, Ben M. Harvey,  
Bas Rokers, Serge O. Dumoulin

### Correspondence

rokers@wisc.edu

### In Brief

Barendregt et al. provide insight into the representation of our visual world in the human brain. Using binocular stimuli and high-field (7 T) MRI, they show that a transformation occurs from a representation based on the two retinal images to a representation based on a unified single (cyclopean) image of the visual world.



# Transformation from a Retinal to a Cyclopean Representation in Human Visual Cortex

Martijn Barendregt,<sup>1,2</sup> Ben M. Harvey,<sup>1,3</sup> Bas Rokers,<sup>1,2,4,\*</sup> and Serge O. Dumoulin<sup>1,4</sup>

<sup>1</sup>Experimental Psychology, Helmholtz Institute, Utrecht University, 3584 CS Utrecht, the Netherlands

<sup>2</sup>Department of Psychology, University of Wisconsin–Madison, 1202 West Johnson Street, Madison, WI 53706, USA

<sup>3</sup>Faculty of Psychology and Education Sciences, University of Coimbra, 3001-802 Coimbra, Portugal

<sup>4</sup>Co-senior author

\*Correspondence: [rokers@wisc.edu](mailto:rokers@wisc.edu)

<http://dx.doi.org/10.1016/j.cub.2015.06.003>

## SUMMARY

We experience our visual world as seen from a single viewpoint, even though our two eyes receive slightly different images. One role of the visual system is to combine the two retinal images into a single representation of the visual field, sometimes called the cyclopean image [1]. Conventional terminology, i.e. retinotopy, implies that the topographic organization of visual areas is maintained throughout visual cortex [2]. However, following the hypothesis that a transformation occurs from a representation of the two retinal images (retinotopy) to a representation of a single cyclopean image (cyclopotopy), we set out to identify the stage in visual processing at which this transformation occurs in the human brain. Using binocular stimuli, population receptive field mapping (pRF), and ultra-high-field (7 T) fMRI, we find that responses in striate cortex (V1) best reflect stimulus position in the two retinal images. In extrastriate cortex (from V2 to LO), on the other hand, responses better reflect stimulus position in the cyclopean image. These results pinpoint the location of the transformation from a retinal to a cyclopean representation and contribute to an understanding of the transition from sensory to perceptual stimulus space in the human brain.

## RESULTS

Our eyes typically receive two slightly different images of the same visual world (Figure 1). In this study, observers viewed a contrast-defined bar stimulus with slight opposite horizontal offsets in each eye (binocular disparity), so that it was perceived behind the fixation plane (Figure 2, “position in depth”). This stimulus moved across the visual field, producing systematic changes in fMRI response throughout early visual cortex [3, 4]. Position in the retinal versus cyclopean image of such a stimulus necessarily differs, such that the cyclopean position is always in between the positions in the two retinal images. We hypothesized that the neural response in visual cortex could reflect either the superimposed positions of the stimuli on both retinas (retinotopic representation [2]) or the single position in the cyclopean image (cyclopotopic representation).

We used two control stimuli to generate two alternative models of the blood-oxygen-level dependent (BOLD) response, one based on stimulus positions in the two retinal images and one based on stimulus position in the cyclopean image. The first control stimulus matched the position of the experimental stimulus in the cyclopean image: a contrast-defined bar was presented at identical locations in each retinal image without binocular disparity (Figure 2, “single position”). The second control stimulus matched the position of the experimental stimulus in the two retinal images without being integrated into a cyclopean image: a contrast-defined bar temporally alternated between the left and right eye images of the experimental stimulus (Figure 2, “offset positions”), so that it stimulated the same retinal locations but due to the temporal alternation was not integrated into a single cyclopean image.

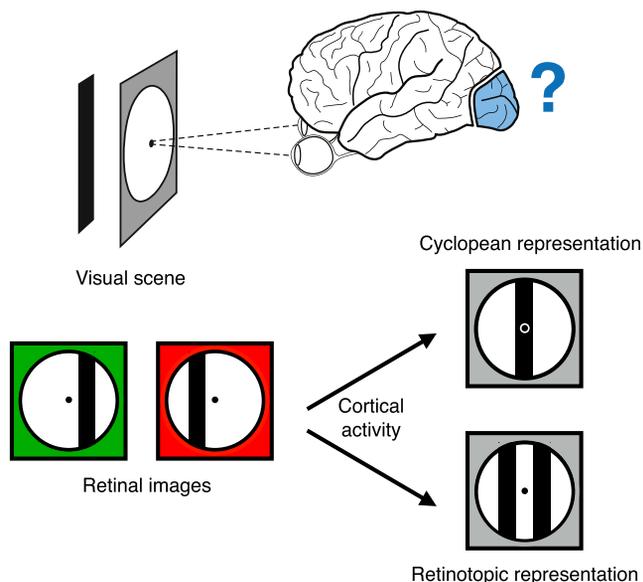
We used the BOLD response to the control stimulus to estimate the population receptive field (pRF), i.e., the region of visual space that optimally stimulated the neural population at each recording site (voxel) [5, 6]. In all experiments, we verified eye vergence by asking observers to perform a demanding task at fixation (on average 77% correct; Figure S1).

## Representations Can Be Discriminated in Early Visual Cortex

We first verified that the two models predicted significantly different responses and that we could resolve differential activity between the two control stimuli using cross-validation. Figure 3A compares how well predictions based on the “single position” and “offset positions” stimuli explained the measured fMRI responses in V1, V2 and V3 for the two control stimulus conditions. This analysis always included all responsive recording sites in each visual area. In V1 and V2, we could reliably discriminate responses elicited by the “single position” and “offset positions” stimuli. In V3, response predictions did not differ sufficiently to identify the stimulus representation when we include all responsive recording sites. Therefore, we limited our first analysis to V1 and V2.

## V1 Represents Retinal Stimulus Location, While Extrastriate Cortex Represents Cyclopean Stimulus Location

To investigate the representation of the “position in depth” stimulus, we compared whether the responses elicited by viewing of the “position in depth” stimulus were better predicted by the pRF models based on either the retinal images or the cyclopean



**Figure 1. Does Cortical Representation of Binocular Stimuli Reflect the Retinal or the Cyclopean Image(s)?**

Elements of a 3D visual scene produce different retinal images in the two eyes. The visual system combines the retinal images into a single cyclopean image of the visual field. We investigated whether the cortical representation of such stimuli reflects their retinal or cyclopean image.

image (Figure 3B). The negative difference in V1 (striped bar,  $t = -3.66$ ,  $p = 2 \times 10^{-5}$ ,  $n \geq 1,135$  voxels/observer) indicates that a prediction based on the retinal model best describes V1's responses. The positive difference in V2 (black bar,  $t = 3.64$ ,  $p = 2 \times 10^{-5}$ ,  $n \geq 1,010$  voxels/observer) indicates that a prediction based on the cyclopean model best describes V2's responses. The difference between the results of V1 and V2 is also significant ( $t = 11.47$ ,  $p < 1 \times 10^{-7}$ ), further demonstrating that these areas have distinct representations of the visual input. Results for each individual observer can be found in Figure S4.

To investigate the representation of position in extrastriate cortex beyond V2, we repeated this analysis using only recording sites that could correctly discriminate between the control conditions (Figure 3B, right panel). By selecting recording sites based on control conditions, we avoid biases when evaluating the “position in depth” condition. We again find that the neural

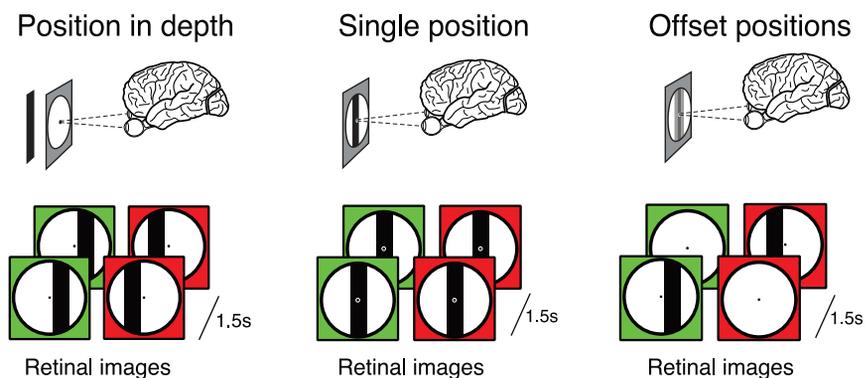
response is best predicted by the retinal model in V1 (striped bar,  $t = -4.03$ ,  $p = 8 \times 10^{-7}$ ,  $n \geq 365$  voxels/observer) and is best predicted by the cyclopean model in V2 (first black bar,  $t = 3.24$ ,  $p < 0.01$ ,  $n \geq 293$  voxels/observer). Furthermore, the responses to the stimulus in V3, V3A, LO1, and LO2 are also best predicted by the cyclopean model (all other black bars,  $t > 2.91$ ,  $p < 0.005$ ,  $n \geq 174$  voxels/observer). As such, the BOLD response throughout extrastriate visual cortex is consistent with the representation of binocular stimuli according to their position in the cyclopean image rather than in the retinal images.

### Responses Reflect Differences in Position, Not Stimulus Features

The temporal interleaving of the “offset positions” control stimulus used in the experiment results in a difference in temporal contrast energy. We therefore tested an additional stimulus condition in three observers to determine whether the predictions made by our models accurately reflect the differences in representation of stimulus location rather than other stimulus features. We used a stimulus with a vertical, instead of horizontal, offset between the two eyes. This stimulus cannot be fused into a single perceived position and therefore behaves similarly to our “offset positions” stimulus. However, this stimulus is continuously presented to both eyes, rather than temporally interleaved, such that the stimulus energy is comparable to the “single position” and “position in depth” stimuli. Using the predictions generated by our control stimuli, we tested whether the responses to this “vertical offset” stimulus are best characterized as a response to two offset positions or a response to a single position. We found that the models consistent with a representation of the retinal images were significantly better at explaining the data obtained with this stimulus in both V1 ( $t = -3.47$ ,  $p = 5 \times 10^{-4}$ ,  $n \geq 1,260$  voxels/observer) and V2 ( $t = -7.94$ ,  $p = 2 \times 10^{-14}$ ,  $n \geq 1,010$  voxels/observer), suggesting that our results reflect a true difference in the representation of stimulus location (Figure 3C).

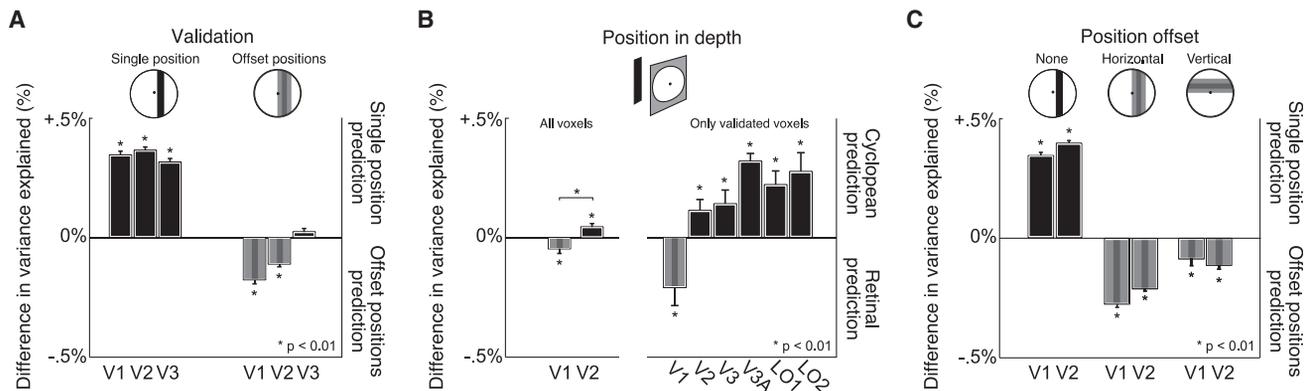
### DISCUSSION

We investigated the cortical responses to binocular stimuli in human visual cortex and found that the neural response best reflects a superimposed representation of the retinal position of visual stimuli in striate cortex (V1). We also found that



**Figure 2. Stimuli Used in the Experiment**

“Position in depth” stimulus: two bars are presented simultaneously to both eyes with a horizontal offset. This results in the percept of a bar positioned in depth so that the stimulated retinal locations differ from the bar's perceived location in the visual field. “Single position” stimulus: a bar is presented in the same position in both eyes and repositioned every 1.5 s (1 TR) to estimate the population receptive field of each cortical location. “Offset positions” stimulus: presentation of the bar is alternated between the two eyes. Bars are presented in the same retinal locations as the “position in depth” stimulus but alternate between the left and right eye to avoid the perception of a bar in depth.



**Figure 3. V1 Responses Reflect Stimulus Position in the Retinal Images; Responses in Extrastriate Cortex Reflect Position in the Cyclopean Image**

(A) In areas V1 and V2, our method successfully distinguishes between a stimulus presented at a single position in the visual field and a stimulus presented at two offset positions on the retina. The difference in variance explained between a model describing single versus offset positions is shown for the two control conditions and early visual areas ( $n = 7$ ). Each bar represents the difference in variance explained between the two models when fitted to data collected using a particular stimulus (as indicated by the stimulus icon above the bars). We demonstrate that data from V1 and V2 correctly distinguish whether “single position” or “offset positions” stimuli were shown. When we include all responsive voxels within each region of interest, we find that area V3 cannot correctly discriminate these two control stimuli for data obtained using the “offset positions” stimulus.

(B) Transformation from a representation of two distinct retinal images to a representation of the cyclopean image across cortical areas. The difference in variance explained between a model describing cyclopean image versus two distinct retinal images is shown for the “position in depth” condition and early visual areas ( $n = 7$ ). The representation of the “position in depth” stimulus in V1 is best explained by the retinal images of the stimulus (striped bar,  $p < 0.01$ ), whereas the representation in V2 is best explained by the cyclopean image (black bar,  $p < 0.01$ ). In the left panel, all responsive voxels in areas V1 and V2 are included in the analysis. The right panel shows the same analysis limited to the voxels in visual areas V1 up to LO2 that are able to correctly discriminate between our two control stimuli. Both the left and right panel show the same transformation between area V1 and V2, going from a representation of the two retinal images in V1 to a representation of the cyclopean image in V2. Furthermore, we see that the responses in subsequent extrastriate areas best reflect the cyclopean image as well.

(C) Responses reflect differences in stimulus position. To exclude the possibility that the differences in responses are due to differences in the stimulus energy of the control stimuli, we tested three observers using a stimulus containing vertical instead of horizontal disparity. The stimulus was presented continuously, so that it had the same stimulus energy as the “single position” stimulus, but unlike the horizontal-disparity stimulus, the vertical-disparity stimulus cannot be integrated into the representation of a single bar in the cyclopean image. We find that the responses to the “vertical offset” stimulus in both V1 and V2 are best described by the retinal position of the stimulus (V1:  $p = 5 \times 10^{-4}$ , V2:  $p = 2 \times 10^{-14}$ ).

All error bars represent mean  $\pm$  SEM.

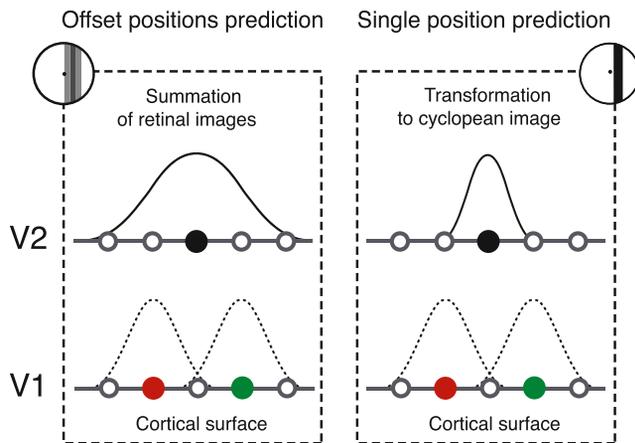
responses in extrastriate cortex (V2 to LO) better reflect the position of visual stimuli in the cyclopean image. Taken together, these results provide evidence for a transformation from two superimposed retinotopic maps to a single cycloptopic map in early visual cortex. Thus, the cortical responses in extrastriate cortex (starting in V2) are to some extent independent of stimulus position in the retinal images and more closely reflect position in the cyclopean image. We would therefore predict that when retinal position of a stimulus is changed (but cyclopean position remains the same), the responses shift across the cortical surface in striate cortex but not extrastriate cortex. For example, two objects that are positioned directly behind each other with respect to an observer will result in responses in different locations in striate cortex but the same location in extrastriate cortex.

These fMRI results in humans are consistent with previous work in other primates showing that in V1, neurons are commonly monocular, preferentially responding to input from one or the other eye, whereas from V2 onward, neurons are mostly binocular, responding to input from either eye [7]. We show here that this change in neural responsivity from V1 to V2 is accompanied by a change in the topographic organization of these visual areas. We would like to emphasize that our results do not provide evidence in favor of a spatiotopic representation of the visual field independent of fixation [11, 12]. Rather, our re-

sults are in line with the notion that the representation in extrastriate cortex reflects visual field location relative to fixation, i.e., a cyclopean representation [1].

The nature of the transformation that occurs between V1 and V2 can be appreciated by comparison with a model in which the representation in extrastriate cortex reflects a simple summation of the two retinal images (Figure 4). If the representation in extrastriate areas were the result of a simple summation of V1’s representation of the two retinal images, then the “offset positions” model, predicting a broader response, would be better at predicting the neural responses. This is indeed the case for stimuli with a vertical disparity, which the visual system does not integrate into a single cyclopean image. For stimuli with a horizontal disparity, which are integrated, we find that simple summation does not predict neural responses as well. Instead, we suggest a transformation where the responses to the retinal images are combined with a corresponding binocular disparity. This means that two objects located exactly behind each other but at different distances from an observer elicit responses at different recording sites in V1, but not in V2 or subsequent extrastriate areas.

A well-known feature of the organization in visual cortex is the increase in population receptive field size between visual areas [5]. Although our data also show this increase between visual areas, we emphasize that this does not explain our results. If



**Figure 4. Responses in Extrastriate Cortex Reflect the Cyclopean Image, Independent of the Position in the Retinal Images**

Our results cannot be explained by assuming that the cortical responses in area V2 are based on the average of the stimulus positions in the two retinal images. Each panel illustrates a schematic view of the cortical surface in V1 and V2. The left panel depicts the prediction of a simple summation of the retinal images underlying the cortical responses in V2. The right panel shows the prediction of a response in V2 that is a transformation into a different position from the positions in the retinal images. Since our “single position” model provides a better prediction of the observed responses in V2, our data suggest that a transformation occurs between V1 and V2, such that responses in V1 reflect stimulus position in the two retinal images, whereas responses in V2 reflect stimulus position in the single cyclopean image.

the difference in responses were due solely to larger population receptive fields in extrastriate cortex, we should not have been able to dissociate between our two control stimuli (single versus offset bars). Given that we can dissociate between these stimuli, we conclude that the increase in population receptive field size is not a sufficient explanation of our findings. Furthermore, the change in the proportion of neurons responding to monocular or binocular input between striate and extrastriate cortex is not sufficient to explain our results [8–10], as binocular input alone does not a priori distinguish between a summed retinal and cyclopean visual field map representation (Figure 4). The goal of this study was not to identify where binocular combination occurs in visual cortex but rather to investigate how the human visual system infers the cyclopean representation from the two retinal images.

Our findings provide neurophysiological evidence of a representation of the cyclopean image in early visual cortex. Further research based on our results could investigate the role of monocular occluded objects in the cyclopean visual space, which has previously been studied only using psychophysical methods [13, 14]. Our results might shed light on psychophysical results showing that many low-level adaptation effects (assumed to be mediated in V1) do not transfer or only partially transfer interocularly, whereas some higher-level effects transfer almost completely.

The topographic organization of sensory and motor cortices is a central pillar of neuroscience and has been instrumental to our understanding of cortical representation. Recently, we demonstrated a topographic organization that does not reflect the layout of sensory organs [15]. In visual cortex, the topographic

maps are commonly thought to mirror the layout of the retina, and hence topographic regions of visual cortex are often referred to as retinotopic. Here, we demonstrate that early visual cortex deviates from the sensory organ layout by representing the cyclopean image instead, which may be better suited to inform subsequent behavior. In conclusion, our results show that the representation of position is systematically transformed between striate and extrastriate cortex. The transformation alters a retinotopic to a cyclopotopic representation. These results contribute to a growing body of work [16, 17] dispelling the notion that V2 is simply a more complex version of V1.

## EXPERIMENTAL PROCEDURES

### Participants

fMRI data were collected from seven participants (one female, aged 27–38; three participants were authors), all with normal or corrected-to-normal visual acuity. All were experienced psychophysical participants in both motion and depth experiments. We tested all participants on their stereo vision (using the same fixation dot task as in the experiment) prior to the experiment and found no abnormalities. The non-author participants were naive to the purpose of the experiment. Experiments were approved by the Medical Ethics Committee of the University Medical Center Utrecht, undertaken with the written consent of each participant, and performed in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

### Stimuli

Visual stimuli were back-projected onto a screen (15 × 7.9 cm) located inside the MRI bore. The participant viewed the display through custom-built prisms that allowed separate presentation for each eye. The total viewing distance from the participant to the display screen was 41 cm.

Stimuli were generated in MATLAB using Psychophysics Toolbox Version 3 [18, 19]. All stimuli were presented in a central, circular aperture with a diameter of 6°. A small dot (0.1°) presented in the center of the aperture was used to maintain fixation. Outside the stimulus aperture, the same pink noise (1/f) background was presented to both eyes, facilitating binocular fusion. See Figure 2 for illustrations of the different stimuli described here.

All stimuli consisted of a moving bar aperture (Figure 2) that moved through the visual field in eight different directions. These bar apertures revealed a random 1/f noise (pink noise) pattern that was re-generated at 10 Hz. The bar had a width of 0.75° and moved through the aperture in 20 discrete steps of 0.3°, each lasting 1.5 s, the repetition time (TR). As such, each bar pass lasted a total of 30 s.

In the main stimulus condition (“position in depth”), binocular disparity was introduced by displacing the bar stimulus by 0.25° in opposite horizontal directions in the two eyes. Perceptually, the stimulus simulated an uncrossed disparity, i.e., the bar is perceived as floating in depth behind the rest of the display.

For the first control stimulus (“single position”), the bars were presented in the same position in both eyes. Because there is no binocular disparity, this stimulus is perceived as a single bar at the same distance as the rest of the display. In the second control stimulus (“offset positions”), the bars were offset between the eyes by the same amount as the main stimulus but were presented alternating between the two eyes, i.e., temporally interleaved, at 10 Hz. This stimulated the same retinal positions as the “position in depth” condition, but without the perceptual experience of a bar floating in depth.

### Fixation Task and Monitoring Binocular Fusion

During the presentation of the stimuli, the participant performed a simple task at fixation. At random intervals, the fixation dot would move either slightly toward or slightly away from the participant in depth. The participant reported the direction of the change (i.e., toward or away) by pressing one of two response buttons. This task ensured fixation of gaze at the center of the display and ensured proper binocular fusion. The difficulty of the task, i.e., the amount of disparity, was adjusted so that most participants would be correct around 75% of the time. The behavioral results for this task (Figure S1) show that participants performed this discrimination task correctly on an average of 77% of trials.

### Magnetic Resonance Imaging

All MRI data were collected at the University Medical Centre Utrecht using a Philips 7 T MRI scanner. T1-weighted anatomical MRI data were acquired using a 32-channel head coil at a resolution of  $0.5 \times 0.5 \times 0.8$  mm. These were subsequently resampled to 1 mm isotropic resolution. Repetition time (TR) was 7 ms, echo time (TE) was 2.84 ms, and flip angle was  $8^\circ$ .

For four participants, functional T2\*-weighted 2D echo planar images were acquired using a 16-channel head coil. For the other three participants, functional images were acquired using a 32-channel head coil. In all cases, the resolution was  $1.98 \times 1.98 \times 2$  mm, field of view was  $190 \times 190 \times 52$  mm, TR was 1,500 ms, TE was 25 ms, and flip angle was  $80^\circ$ . The acquired volume was always oriented perpendicular to the calcarine sulcus, providing coverage of the occipital lobe and posterior parts of the parietal and temporal lobes.

Functional runs were each 248 time frames (372 s) in duration, of which the first eight time frames (12 s) were discarded to ensure the signal was at steady state. Data for all conditions were collected during 2–3 sessions per participant, for a total of 7–9 repetitions per condition.

### Processing of Functional Imaging Scans

Functional scans were first compensated for head movement and motion artifacts [20]. Subsequently, the functional images were averaged and aligned to the whole-brain anatomical scan. The alignment was performed automatically [20] and afterward checked and refined manually if needed.

### Model-Based fMRI Analysis

The population receptive field (pRF) is defined as the region of visual space that optimally stimulates a recording site [5]. Using a previously described method, we estimated pRF properties (position and size) as well as the hemodynamic response function (HRF) from the fMRI data [5, 6].

Briefly, the pRF method is based on a forward model that estimates the pRF position and size based on the time course of the stimulus aperture and the measured BOLD time series. As illustrated in Figure S2, we model the pRF as a 2D difference of Gaussians function described by four parameters (position:  $x, y$ ; size:  $\sigma_{\text{center}}, \sigma_{\text{surround}}$ ). We multiply each candidate pRF with the stimulus aperture at each point in time, resulting in a predicted time course of the neural activation. After convolution with the HRF, this yields a prediction of the BOLD response, given the candidate pRF. Next, the predicted BOLD response is compared to the measured BOLD time series, and the residual sum of squares (RSS) is used to assess the goodness of fit. We use a coarse-to-fine procedure to identify the optimal pRF parameters. We start with a large set of permutations of possible pRF parameters. For each recording site, the optimal parameters from this large set are refined using a nonlinear optimization routine. The pRF parameters that produce the prediction with the smallest RSS are chosen for each recording site.

The analysis in this study differs from the typical pRF analysis because we use two different stimulus apertures to make the candidate time series predictions. This results in two different predictions for each recording site's time series. As an additional step at the end of the analysis, we compare which of the predicted time series better accounts for the measured BOLD time series across recording sites in terms of the amount of variance explained. Figure S3 shows an example BOLD time series with the pRF model predictions for a recording site in area V1 of a single participant. As shown in the top panel (black dots: measured time series, blue line: single-bar prediction, red line: offset-bars prediction), the differences are quite small since the difference between the two stimulus sequences used to generate the predictions is small. However, looking at the difference between the two pRF predictions (bottom panel, green line), it is evident that the differences are mainly located at the points in the time series where the bar makes a horizontal (left to right, v.v.) or diagonal pass through the visual field. This is what we would expect, since the binocular disparity is only present when the stimulus passes through the visual field in horizontal (left to right, v.v.) and diagonal directions.

### Selecting Voxels

In our main analysis, we included all voxels whose best-fitting pRF model could explain more than 25% of the observed BOLD response variance (in addition to criteria such as eccentricity and size of the pRF) to exclude unre-

sponsive voxels. We found, however, that many such voxels in areas beyond V2 could not reliably discriminate between the two control stimuli used in the experiment, although they still explained over 25% of response variance. To investigate cortical representations in these areas, we ran an additional analysis that used only voxels whose pRF models in our two control conditions could correctly discriminate which stimulus was shown. In order to be included in the analysis, a voxel had to correctly discriminate between the two pRF models based on the original time series as well as explain at least 25% of the variance in the time series.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2015.06.003>.

### ACKNOWLEDGMENTS

This work was supported by Netherlands Organization for Scientific Research (NWO) Onderzoekstalent Grant 406-11-197 to M.B. and F.A.J. Verstraten, NWO Veni Grant 451-09-030 to B.R., and NWO Vidi Grant 452-08-008 to S.O.D.

Received: February 5, 2015

Revised: May 13, 2015

Accepted: June 1, 2015

Published: July 2, 2015

### REFERENCES

- Julesz, B. (1971). *Foundations of Cyclopean Perception*. (University of Chicago Press).
- Inouye, T. (1909). *Die Sehstörungen bei Schussverletzungen der kortikalen Sehphäre*. (Engelmann).
- Engel, S.A., Rumelhart, D.E., Wandell, B.A., Lee, A.T., Glover, G.H., Chichilnisky, E.J., and Shadlen, M.N. (1994). fMRI of human visual cortex. *Nature* 369, 525.
- Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268, 889–893.
- Dumoulin, S.O., and Wandell, B.A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647–660.
- Zuiderbaan, W., Harvey, B.M., and Dumoulin, S.O. (2012). Modeling center-surround configurations in population receptive fields using fMRI. *J. Vis.* 12, 10.
- Chen, G., Lu, H.D., and Roe, A.W. (2008). A map for horizontal disparity in monkey V2. *Neuron* 58, 442–450.
- Talbot, S.A., and Marshall, W.H. (1941). Physiological studies on neural mechanisms of localization and discrimination. *Am. J. Ophthalmol.* 24, 1255–1264.
- Hayhow, W.R. (1958). The cytoarchitecture of the lateral geniculate body in the cat in relation to the distribution of crossed and uncrossed optic fibers. *J. Comp. Neurol.* 110, 1–63.
- Blasdel, G.G., and Fitzpatrick, D. (1984). Physiological organization of layer 4 in macaque striate cortex. *J. Neurosci.* 4, 880–895.
- d'Avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D.C., and Morrone, M.C. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nat. Neurosci.* 10, 249–255.
- Gardner, J.L., Merriam, E.P., Movshon, J.A., and Heeger, D.J. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatio-topic. *J. Neurosci.* 28, 3988–3999.
- Erkelen, C.J., and van Ee, R. (2002). The role of the cyclopean eye in vision: sometimes inappropriate, always irrelevant. *Vision Res.* 42, 1157–1163.

14. Ono, H., Mapp, A.P., and Howard, I.P. (2002). The cyclopean eye in vision: the new and old data continue to hit you right between the eyes. *Vision Res.* *42*, 1307–1324.
15. Harvey, B.M., Klein, B.P., Petridou, N., and Dumoulin, S.O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science* *341*, 1123–1126.
16. Tong, F., and Engel, S.A. (2001). Interocular rivalry revealed in the human cortical blind-spot representation. *Nature* *411*, 195–199.
17. Freeman, J., Ziemba, C.M., Heeger, D.J., Simoncelli, E.P., and Movshon, J.A. (2013). A functional and perceptual signature of the second visual area in primates. *Nat. Neurosci.* *16*, 974–981.
18. Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat. Vis.* *10*, 433–436.
19. Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* *10*, 437–442.
20. Nestares, O., and Heeger, D.J. (2000). Robust multiresolution alignment of MRI brain volumes. *Magn. Reson. Med.* *43*, 705–715.

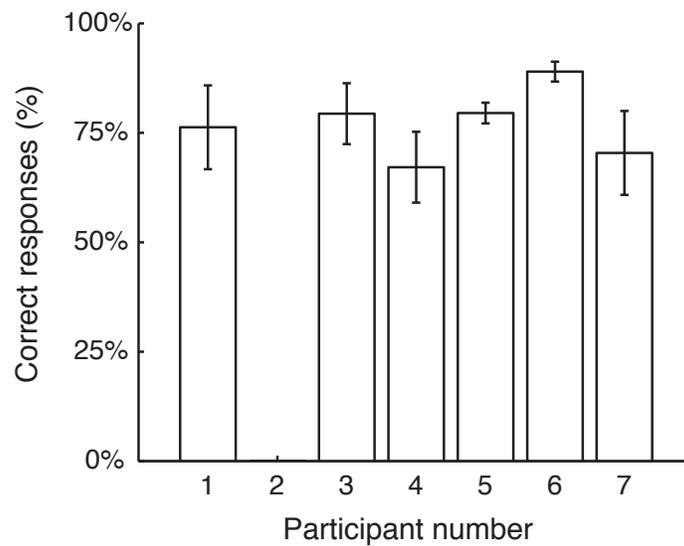
**Current Biology**

**Supplemental Information**

## **Transformation from a Retinal to a Cyclopean**

## **Representation in Human Visual Cortex**

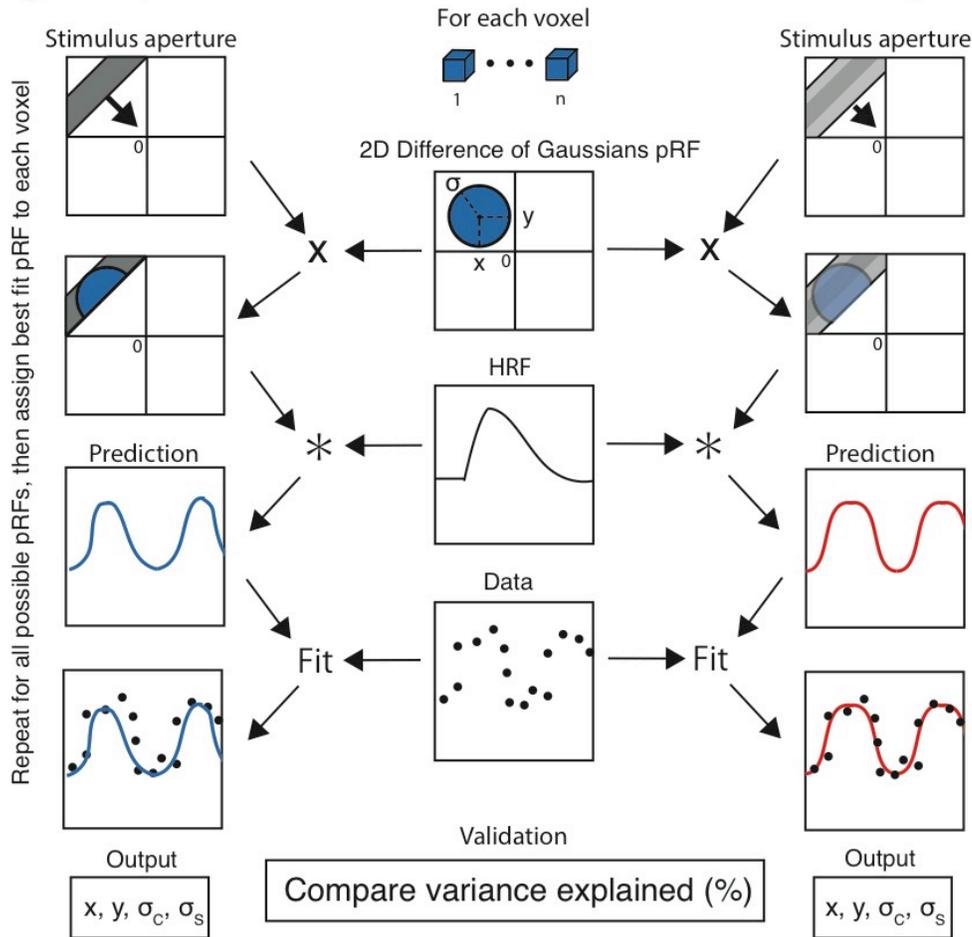
**Martijn Barendregt, Ben M. Harvey, Bas Rokers, and Serge O. Dumoulin**



**Figure S1 (related to Figure 2) Behavioral results for all participants.** At random intervals during the experiment the fixation dot's position in depth was changed briefly. Participants reported the direction of the change in position (towards or away) by pressing one of two buttons on the response box. The results are plotted as the percentage of correct responses during the experiment for each participant. Data for participant 2 is missing due to a technical problem with the response box. All error bars  $\pm$  s.e.m.

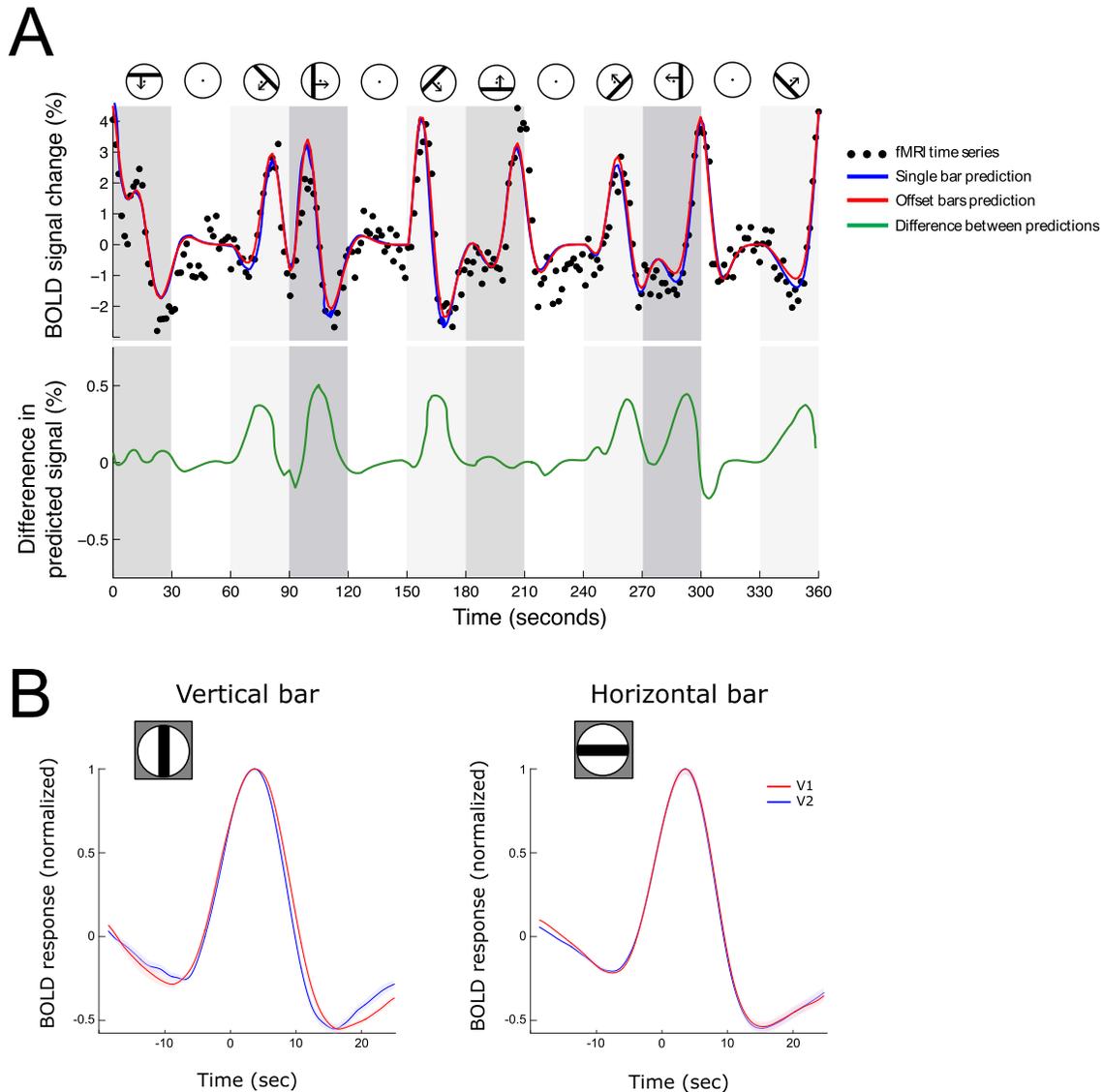
### Single bar prediction

### Offset bars prediction

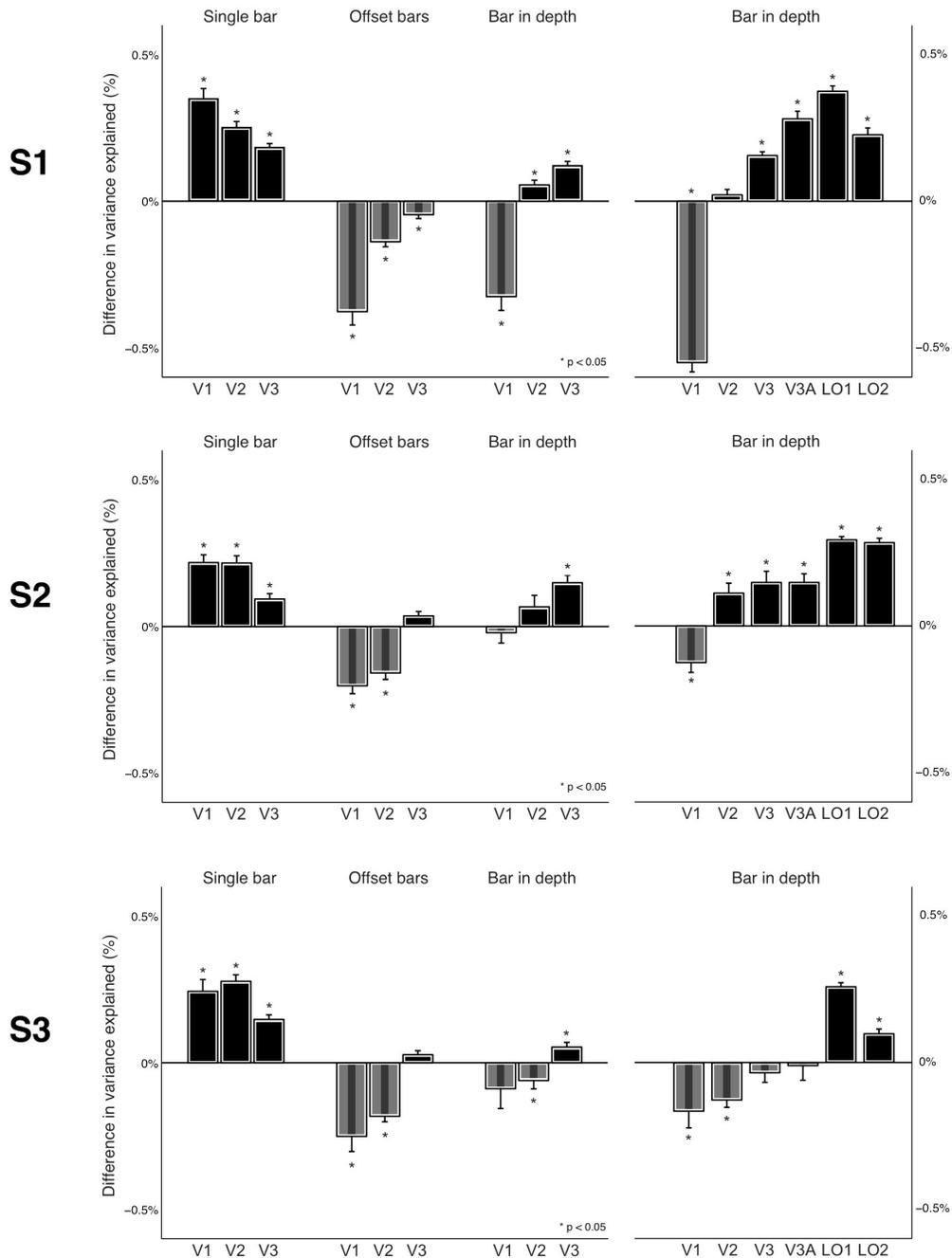


**Figure S2 (related to Figure 2) Schematic description of the pRF modeling analysis.**

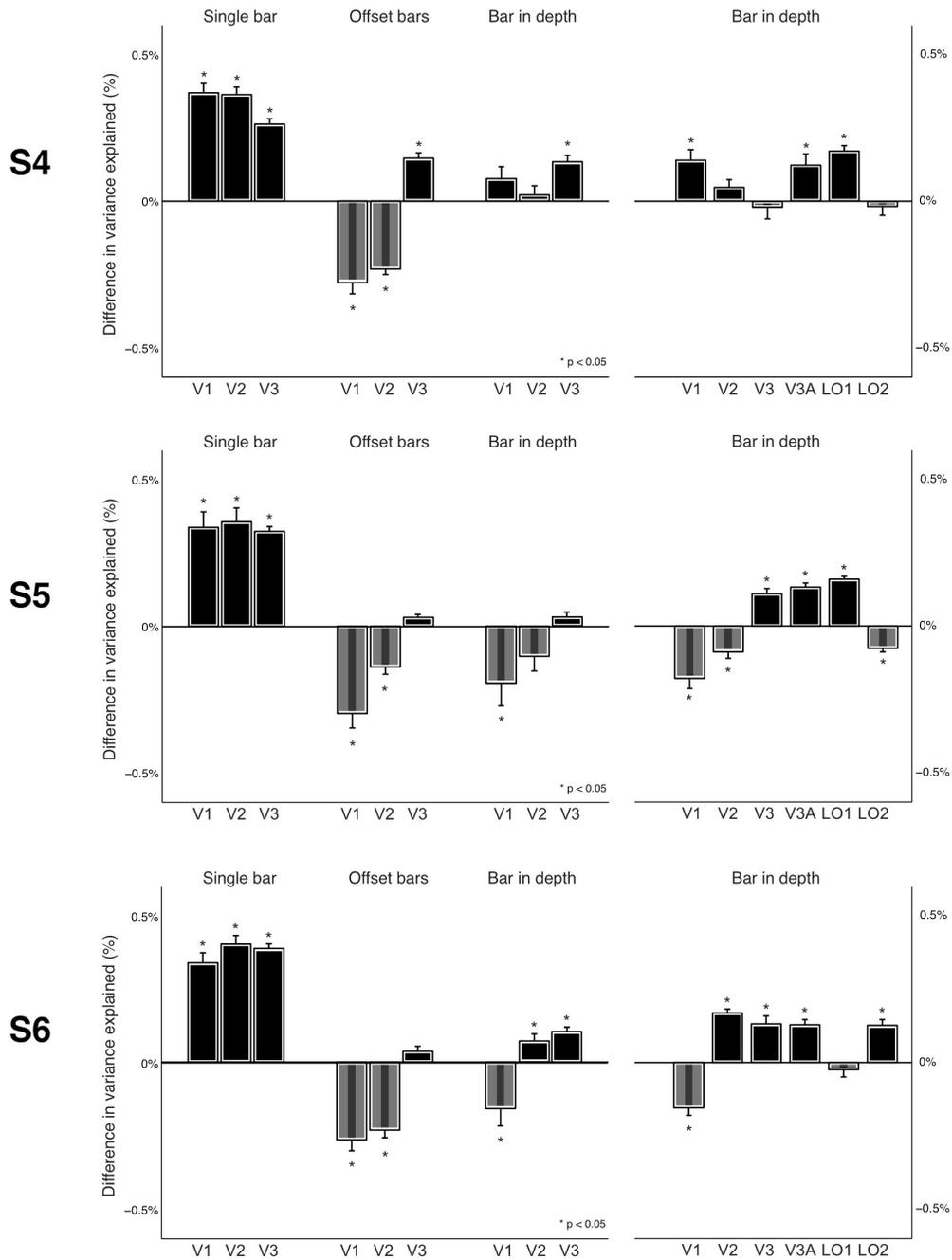
Flowchart describing the pRF analysis for a single MRI recording site. We computed the overlap of the stimulus aperture with a model of the pRF (modeled as a 2-D Difference of Gaussians) for a given recording site and convolved the resulting time series with the haemodynamic response function (HRF) to provide a prediction of the measured time series for each recording site. Using different parameters for the modeled pRF we searched for the best fitting prediction to the measured data. When applying this method using different stimulus apertures, we can establish which model better predicts the observed data. The models are compared by the amount of variance explained in the observed BOLD time series.



**Figure S3 (related to Figure 4) A: Examples of an observed V1 recording unit (voxel) time series and the predicted pRF model time series.** pRF model predictions generated using two different stimulus apertures (single bar and offset bars) look relatively similar by eye and both predict a large proportion of the variance in the fMRI time series (panel 1). Panel 2 reveals that the major differences in the predicted time courses coincide with vertical and diagonal orientations of the bar stimuli. This is expected since the retinal positions of the stimulus in the two eyes differ for diagonal and vertical orientations of the stimulus, whereas no such difference in retinal position exists for the horizontal stimulus orientation. **B: The observed HRF response averaged across voxels, stimulus repetitions and observers.** The normalized BOLD response (vertical axis) is plotted as a function of time relative to the center of the pRF (horizontal axis). Because differences in pRF position between voxels result in different time courses we used a previously described method to align the BOLD response to the same point in time: the moment the stimulus passes through the center of each pRF (*Dumoulin, Hess, May, Harvey, Rokers & Barendregt (2014) Journal of Vision, 14(5):18*). We observe a slightly broader HRF in visual area V1 compared to visual area V2 when the position-in-depth stimulus is oriented vertically, but not when it is oriented horizontally.

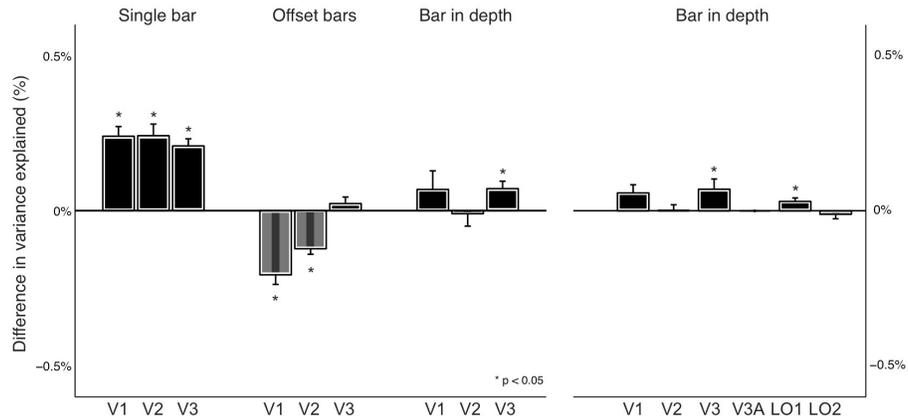


**Figure S4 (related to Figure 3) Individual participant results for participants S1-S3.** Difference in variance explained between a model encoding a single stimulus position and a model encoding two offset retinal positions for the 3 datasets and three early visual areas. The solid black bars indicate that the data is best explained by a model encoding a single stimulus position and the striped bars indicate the data is best explained by a model encoding two offset retinal stimulus positions. As with **Figure 3** in the main text, the right panel shows the difference in variance explained only for voxels that could correctly discriminate between the stimuli. All error bars  $\pm$  s.e.m.



**Figure S4 (cont.) Individual participant results for participants S4-S6.** Difference in variance explained between a model encoding a single stimulus position and a model encoding two offset retinal positions for the 3 datasets and three early visual areas. The solid black bars indicate that the data is best explained by a model encoding a single stimulus position and the striped bars indicate the data is best explained by a model encoding two offset retinal stimulus positions. As with **Figure 3** in the main text, the right panel shows the difference in variance explained only for voxels that could correctly discriminate between the stimuli. All error bars  $\pm$  s.e.m. **Word did not find any entries for your table of contents.**

S7



**Figure S4 (cont.) Individual participant results for participant S7.** Difference in variance explained between a model encoding a single stimulus position and a model encoding two offset retinal positions for the 3 datasets and three early visual areas. The solid black bars indicate that the data is best explained by a model encoding a single stimulus position and the striped bars indicate the data is best explained by a model encoding two offset retinal stimulus positions. As with **Figure 3** in the main text, the right panel shows the difference in variance explained only for voxels that could correctly discriminate between the stimuli. All error bars  $\pm$  s.e.m.