NeuroImage 215 (2020) 116780

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

fMRI representational similarity analysis reveals graded preferences for chromatic and achromatic stimulus contrast across human visual cortex

Erin Goddard, Kathy T. Mullen *

McGill Vision Research, Department of Ophthalmology & Visual Sciences, McGill University, Montreal, QC, H3G1A4, Canada

ARTICLE INFO

Keywords:
fMRI
Visual cortex
Colour/color vision
hV4
VO
Representational similarity analysis (RSA)

ABSTRACT

Human visual cortex is partitioned into different functional areas that, from lower to higher, become increasingly selective and responsive to complex feature dimensions. Here we use a Representational Similarity Analysis (RSA) of fMRI-BOLD signals to make quantitative comparisons across LGN and multiple visual areas of the low-level stimulus information encoded in the patterns of voxel responses. Our stimulus set was picked to target the four functionally distinct subcortical channels that input visual cortex from the LGN: two achromatic sinewave stimuli that favor the responses of the high-temporal magnocellular and high-spatial parvocellular pathways, respectively, and two chromatic stimuli isolating the L/M-cone opponent and S-cone opponent pathways, respectively. Each stimulus type had three spatial extents to sample both foveal and para-central visual field. With the RSA, we compare quantitatively the response specializations for individual stimuli and combinations of stimuli in each area and how these change across visual cortex. First, our results replicate the known response preferences for motion/flicker in the dorsal visual areas. In addition, we identify two distinct gradients along the ventral visual stream. In the early visual areas (V1–V3), the strongest differential representation is for the achromatic high spatial frequency stimuli, suitable for form vision, and a very weak differentiation of chromatic versus achromatic contrast. Emerging in ventral occipital areas (V4, VO1 and VO2), however, is an increasingly strong separation of the responses to chromatic versus achromatic contrast and a decline in the high spatial frequency representation. These gradients provide new insight into how visual information is transformed across the visual cortex.

1. Introduction

A fundamental property of the visual system is the emergence of functional specializations between the lower and higher visual areas. Early cortex receives information from the distinctive magno-, parvo- and koniocellular subdivisions of the lateral geniculate nucleus (LGN), and distinct cortical areas transform and combine these inputs to form our perceptual representations, including motion, stereo, form, and object properties such as colour. fMRI has proved useful for revealing and localizing these specialized cortical functions. The simplest approach is to compare amplitudes of BOLD responses to different stimuli. This reveals areas with response biases to stimuli targeting a particular specialization, such as area hMT for moving stimuli and area FFA for faces (Kanwisher et al., 1997; Tootell et al., 1995). It is limited, however, to revealing univariate biases, since signals are averaged within a region or subregion. In contrast, classification analyses are more sensitive, detecting multivariate signals that reliably vary with the stimulus dimension of interest (e.g. Haxby et al., 2014; Kamitani and Tong, 2005). However, classifier accuracy depends upon multiple unrelated factors that vary across regions, including the number of voxels and BOLD signal strength, making comparisons across regions difficult to interpret.

Moreover, especially in the early visual areas, neural responses contain multiplexed information about multiple stimulus dimensions, making 'stimulus specialization' a concept of limited usefulness when seeking to characterize visual cortex. For example, if a region's response can be used to decode stimulus orientation, colour, and spatio-temporal frequency, then is it useful to consider that the region is specialized for any or all of these stimulus features? Seeking a more holistic view of inter-area biases and specializations, here we identify the evolution of stimulus representations across different regions when each contains information about multiple stimulus dimensions. We use Representational Similarity Analysis (RSA) to facilitate quantitative comparisons across visual areas. RSA, like classifier accuracy, will be adversely affected by low signal strength, but since responses in each region are correlated with multiple models the relative performance of these models can be compared across ROIs.

* Corresponding author.
E-mail address: kathy.mullen@mcgill.ca (K.T. Mullen).

https://doi.org/10.1016/j.neuroimage.2020.116780
Received 21 November 2019; Received in revised form 18 March 2020; Accepted 24 March 2020
Available online 8 April 2020

RSA was first developed in vision for the understanding of object processing (Kriegeskorte et al., 2008), and has been applied in a range of predominantly high-level visual and cognitive functions, including memory, semantics and emotion (Borghesani et al., 2016; Skerry and Saxe, 2015), but has been little used for lower-level vision or colour vision (Bannert and Bartels, 2017, 2013; Bird et al., 2014; Goddard et al., 2017; Goddard and Mullen, 2019; Salmela et al., 2016; Wardle et al., 2016). Furthermore, in the visual object literature, it is increasingly clear that stimulus differences along lower-level feature dimensions may contribute to effects that have been attributed to 'high-level' feature coding (Andrews et al., 2015). In this way, understanding inter-area differences in lower-level feature coding not only provides insight into how cortical areas combine and transform subcortical input, but also provides a valuable frame of reference for investigating the emergence of higher-level coding.

Here we employed an RSA to make quantitative comparisons across LGN and visual cortical areas of the low-level stimulus information encoded in the patterns of voxel responses. Our stimulus set was selected to bias responses towards each of the four functionally distinct subcortical channels that input visual cortex from the LGN. We used two chromatic and two achromatic sinewave stimuli, each of three spatial extents. The chromatic stimuli isolated the L/M-cone-opponent and S-cone-opponent pathways of the LGN, representing the chromatic processing of the parvo- and koniocellular pathways, respectively (Dacey, 2000; Martin and Lee, 2014). The two achromatic stimuli were chosen on the basis of primate lesion data as being biased towards the functioning of either the parvocellular LGN (high spatial, low temporal frequency) or the magnocellular (low spatial, high temporal frequency) LGN pathways (Merigan et al., 1991; Merigan and Maunsell, 1993, 1990; Schiller et al., 1990). While ideally, we would measure responses to a larger set of stimuli to reflect the entire range of spatiotemporal responses, we chose these four stimuli as diagnostic of the LGN subdivisions that form the input to V1. Our aim was to compare quantitatively the response specializations for these stimuli and combinations of stimuli in each visual area and how these change from the LGN and from lower to higher visual areas. We compared seven possible models of how responses to these individual stimuli or combinations of these stimuli predict classifier performance in each visual area. This allows us to make testable predictions for whether cortical responses are associated with specific subcortical inputs or how they are combined. Further, we aimed to identify whether there are gradients of preference for stimulus features as specializations develop between earlier and higher visual areas in the dorsal and ventral streams.

2. Materials and methods

2.1. Participants

For the psychophysical experiment, we collected data on 10 participants (7 female, 3 male, aged 20–50 years), but data from 2 participants (1 female, 1 male) were excluded since they did not complete all conditions. For the fMRI experiments we collected data on 8 participants (5 female, 3 male, aged 21–35 years), including 4 participants (2 female, 2 male) who also completed the psychophysical experiment. All participants were healthy with no history of neurological and/or psychiatric disorders and provided informed consent. Each participant had normal or corrected-to-normal visual acuity, and normal colour vision as assessed with Ishihara plates (Ishihara, 1990) and the Farnsworth-Munsell 100-hue test (Farnsworth, 1957). Both experiments were approved by the Ethics Review Board of the McGill University Health Centre and were conducted in accordance with the Declaration of Helsinki.

2.2. Visual stimuli

For both experiments the stimuli were radial sinewave gratings, as used in previous work (Mullen et al., 2015, 2007). Stimulus contrast was either achromatic (Ach), isoluminant 'red-green' (RG) or 'blue-yellow' (BY), modulated about a mean grey, isolating the luminance, L/M cone opponent and S-cone mechanisms respectively, as shown in Fig. 1. There were two versions of the achromatic stimulus, with spatiotemporal parameters selected to produce greater stimulation in the magnocellular (Ach M-type) or parvocellular (Ach P-type) pathways. The spatiotemporal parameters were selected on the basis of LGN lesion studies in the macaque (see review by Merigan and Maunsell, 1993), which show that parvocellular lesions dramatically reduce sensitivity to static stimuli of high spatial frequencies (~98% reduction at 5 cycles/deg), whereas magnocellular lesions produce marked loss of sensitivity to high temporal frequencies (~90% reduction at 10 Hz). The RG, BY and Ach M-type stimuli each had a spatial frequency of 0.5 cycles/deg, while the Ach P-type stimulus had a spatial frequency of 5 cycles/deg. The RG, BY and Ach P-type stimuli had a 2 Hz sinusoidal contrast phase alternation, while the Ach M-type stimulus was modulated at 10 Hz. The low spatial frequency of the two chromatic stimuli reduces luminance artifacts generated by chromatic aberration for the chromatic stimuli (Bradley et al., 1992; Cottaris, 2003; Mullen, 1985).

Each stimulus was one of three sizes, either 1.5, 13 or 177 degs^2 visual angle, as shown in Fig. 1, each representing a whole number of spatial cycles, plus a half cycle around the fixation marker, meaning that each stimulus had a smooth transition at its outer edge. Stimuli with spatial frequency of 0.5 cycles/deg comprise a total of 0.5, 1.5 or 5.5 cycles, while stimuli of 5 cycles/deg comprise 5, 15 or 55 cycles. A small fixation marker was displayed in the centre of all stimuli (a black dot). Outside the stimulus area, the screen was at its mean luminance. We included three sizes as a control since there are different retinal biases for these different types of stimuli with some having a representation more confined to central vision (L/M colour, high SF achromatic) than others (Anderson et al., 1991; Mullen and Kingdom, 2002; Vanni et al., 2006). Moreover, there are differences in each visual area's magnification of the fovea (Schira et al., 2009) that we aimed to capture by including a range of stimulus sizes.

Stimulus chromaticity was defined in a three-dimensional cone contrast space, with each axis representing the quantal catch of the L, M and S cone types normalized with respect to the grey background (i.e. cone contrast). The vector direction and length within this space defines chromaticity and cone contrast respectively. We determined isoluminance of the RG stimuli for each subject individually based on perceptual minimum motion settings as previously described (Mullen et al., 2010, 2007). We also verified the angle of the BY mechanism within each participant's isoluminant plane by varying vector angle and selecting the direction of minimum visibility (Michna et al., 2007).

We first measured each participant's stimulus detection thresholds psychophysically (see below). For the fMRI experiment, we chose stimulus contrasts close to the maximum of the monitor gamut for the chromatic directions (4% for RG, 30% for BY, and 50% for Ach). We chose high contrast values to yield stimuli that were highly visible even at the smallest stimulus size, with the aim of driving robust responses in visual cortex. We have previously found that stimuli like these evoked similar amplitudes of BOLD response (Mullen et al., 2010). Based on the psychophysical data described below, the RG contrast values ranged from 8 x detection threshold (smallest stimulus) to 22 x detection threshold (largest stimulus). For the remaining stimuli, contrast values across stimulus size ranged from 3 to 16 x detection threshold (BY), 27-146 x detection threshold (Ach, M-type) and 19-45 x detection threshold (Ach, P-type).

2.3. Display apparatus and calibrations

For all psychophysical experiments, we used MATLAB R2006a in combination with a ViSaGe video graphics board with 14 bits of contrast resolution (Cambridge Research Systems Ltd, Rochester, UK) housed in a Pentium PC and displayed on a CRT monitor (Diamond Pro, 2070). The BOLD screen, LCD projector and the CRT display were each linearized

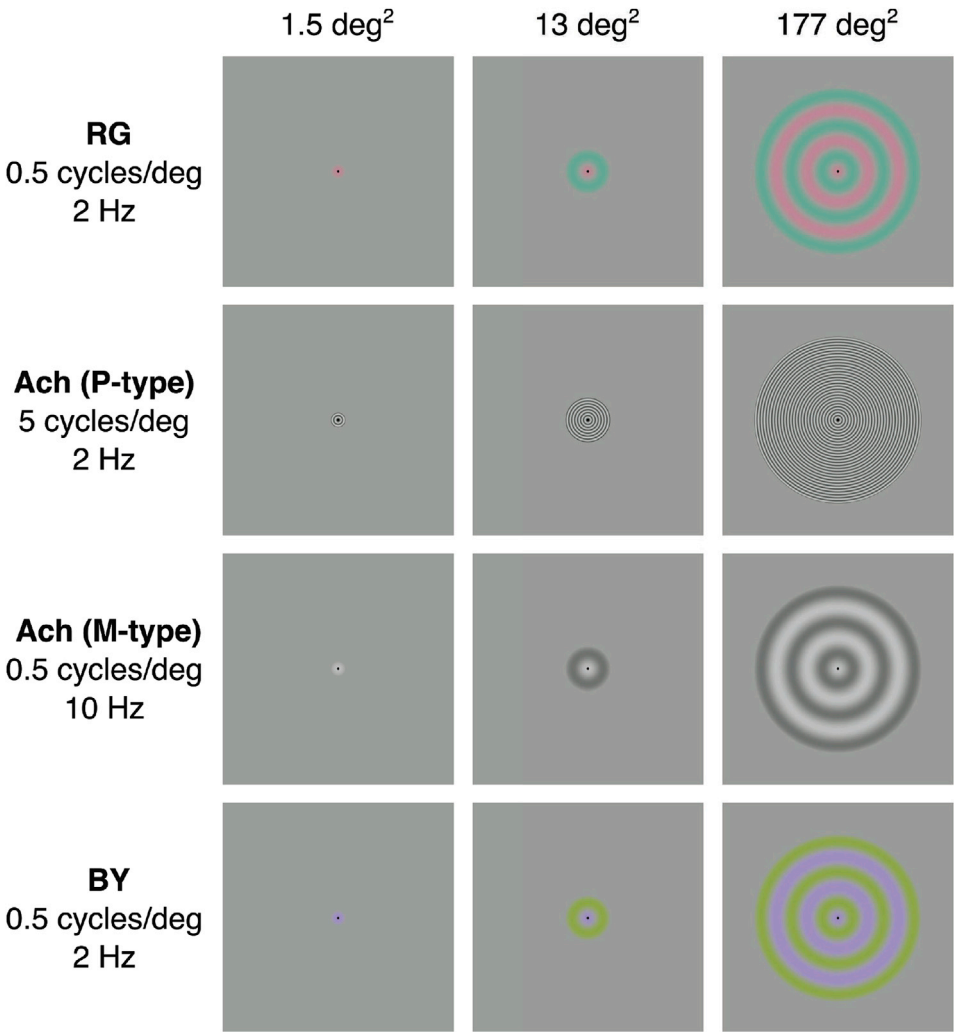

Fig. 1. Visual stimuli. In all experiments, the visual stimuli were radial sinewave gratings isolating the L/M cone opponent (red-green/RG), luminance (Achromatic/Ach) or S-cone (blue-yellow/BY) mechanisms, shown at higher contrast in these illustrations. The achromatic stimuli were either low spatial frequency (0.5 cycles/deg) and high temporal frequency (10 Hz), termed 'M-type' or high spatial frequency (5 cycles/deg) and low temporal frequency (2 Hz) termed 'P-type'. Each stimulus was presented at three sizes, as shown, including a whole number of spatial cycles, plus a half cycle around the fixation marker. For the 0.5 cycles/deg stimuli, stimuli comprised 0.5, 1.5 or 5.5 cycles. For the 5 cycles/deg stimulus (Ach, P-type), stimuli comprised 5, 15 or 55 cycles.

and colour calibrated as described previously (Michna et al., 2007; Mullen et al., 2007).

For fMRI experiments, we displayed stimuli on a 32" BOLD screen LCD monitor (Cambridge Research Systems Ltd, Rochester, UK, resolution 1920x1080). Participants viewed the BOLD screen, which was located at the rear of the MRI bore, through a mirror mounted on the head coil. The total viewing distance was 125 cm. We used a Macbook Pro (2015) running MATLAB (R2017a) in conjunction with routines from Psychtoolbox 3.0 (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) to generate the stimuli and draw them to the BOLD screen (refresh rate 60Hz, mean luminance 52.4 cd/m2).

2.4. Psychophysical methods

Detection thresholds were measured using a standard 2IFC staircase protocol. Interval onsets were separated by 1 s, and the start of each interval was indicated by a tone. The mean grey background and central fixation marker were present through the experiment. On each trial, during one randomly chosen interval, the stimulus appeared within a Gaussian temporal envelope (sigma 125 ms, total duration 500 ms), and the participant's task was to report the interval with a stimulus. The contrast of the test stimulus was varied using a 2-up 1-down staircase procedure: after 2 correct responses at a given contrast, the contrast was lowered by 10%, while after an incorrect response the contrast was increased by 20%. Each staircase was terminated after 6 reversals, with a reversal defined as an incorrect response following a correct response. For each condition we simultaneously acquired data for 2 staircases, with trials randomly interleaved. Each participant completed 2 runs (4 staircases) for each stimulus at each size. The order in which data from the different stimuli were acquired was counter-balanced across participants, but each participant completed every condition once before moving to the repeat. Detection threshold was defined as the average of the contrasts at which reversals occurred.

2.5. fMRI methods: experimental design

For the fMRI experiments, we acquired functional measurements of the BOLD response in a 1-h session for each participant. Each session comprised six runs of 7 min 18 s, which included 29 blocks. Each block commenced with 3 s of the fixation screen (blank grey screen of mean luminance, with fixation dot) followed by 12 s of stimulus. At the end of 29 blocks, there was a final 3 s fixation screen. The 12 s of stimulus included either 4 trials of a contrast discrimination task (described below), or 12 s of a reference stimulus (either a constant black screen or a constant grey screen). The first block was always a grey reference stimulus, and a pair of grey and black reference stimuli occurred in the middle (blocks 14–15) and at the end (blocks 28–29) of each run, with the order of grey and black blocks within these pairs counterbalanced across runs. Each stimulus block included a single stimulus from the 12 unique stimuli (4 stimulus types at 3 sizes) shown in Fig. 1. The 12 unique stimuli were presented in one block each across blocks 2–13 and blocks 16–27 of each run. The order of these stimuli was counterbalanced across runs such that

every stimulus was preceded and followed by every other stimulus once across the session. The order of these counterbalanced runs varied across participants.

During the 12 s stimulus periods, participants performed a contrast discrimination task used in previous work (Goddard et al., 2019; Mullen et al., 2015, 2010, 2007). For each 3 s trial of the task, the ring stimulus was presented twice with a near-threshold contrast difference between them (a 20% contrast increment added to one stimulus and a 20% decrement to the other, yielding a 40% contrast difference about the mean contrast). Each stimulus was presented within a Gaussian temporal envelope (sigma 125 ms, total duration 500 ms), with a 500 ms ISI. In the remaining trial time (1.5 s) the participants indicated with a button press which interval contained the higher contrast stimulus.

2.6. fMRI methods: retinotopic and functional localisers

We identified the visual cortical regions V1, V2, V3, V3A/B, LO1/LO2 and hV4 for each participant using rotating wedge stimuli and expanding and contracting concentric rings (Engel et al., 1994; Sereno et al., 1995), following standard definitions of these areas (Brewer et al., 2005; Goddard et al., 2011; Larsson and Heeger, 2006), and following the model of the foveal confluence presented in Schira et al. (2009). To localize areas VO1 and VO2 we used data from the retinotopic mapping scans in conjunction with a VO localizer, based on this region's preference for chromatic over achromatic contrast (Mullen et al., 2007). To localize hMT + we used a localizer stimulus similar to that described previously (Huk et al., 2002), with 10 s blocks of moving and static dots, interspersed with blank intervals (also 10 s duration). Full details of our retinotopic mapping procedures have been described previously (Goddard et al., 2019).

2.7. fMRI methods: scanning protocols

All magnetic resonance imaging took place at the McConnell Brain Imaging Centre (Montréal, Canada). Functional T2* MR images were acquired on a 3T Siemens MAGNETOM Prisma system with 32-channel head coil. Gradient-echo pulse sequences were used to measure blood oxygenation level-dependent (BOLD) signal as a function of time. We used a scanning protocol with partial head coverage (including the occipital cortex and the LGN, with slices oriented approximately parallel to the calcarine sulcus), no acceleration, and fine spatial resolution (TR = 3000 ms, TE = 38 ms, 28 axial slices, 1.5 mm^3 resolution). Localization of hMT+ was performed in a separate scan with a multiband acceleration factor of 3 (39 axial slices, 1.5 mm^3 resolution, TR = 1210 ms, TE = 30.4 ms). Head movement was limited by foam padding within the head coil. For each participant, we acquired two high-resolution three-dimensional T1 images using an MP-RAGE sequence (TI = 900 ms, TR = 2300 ms, TE = 3.41 ms, 1.0 mm^3 resolution), and averaged these two images to generate the participant's anatomical template.

2.8. fMRI analysis: surface definition and preprocessing of functional data

For each participant's template anatomical, we used the automatic segmentation processes from Freesurfer 6.0 (Dale et al., 1999; Fischl et al., 1999) to define the grey/white matter and pial/grey matter boundaries. For all other MRI data processing, we used AFNI/SUMA packages (AFNI 17.2.12, Sep 6, 2017; Cox, 1996; Saad et al., 2004). All functional data were preprocessed using slice-time correction and rigid-body motion correction before being aligned to the participant's anatomical template. Retinotopic mapping data were then projected onto the cortical surface by averaging between the white and pial boundaries, spatially smoothed (Gaussian filter, full-width at half maximum of 4 mm), and we scaled data for each surface node by the node's mean response across the run. Functional data from the experimental runs were not projected into a surface space, but the volume data were smoothed (Gaussian filter, full-width at half maximum of 3 mm), and data from each voxel scaled by its mean response for the run.

2.9. fMRI analysis: general linear modeling (GLM)

Data collected during the phase-encoded retinotopic mapping scans (rotating wedge and expanding ring) were analyzed with AFNI script *@RetinoProc* (Saad et al., 2001). For all the remaining functional data (VO localizer, hMT + localizer, and experimental data) BOLD responses were modelled using a GLM using AFNI script *3dDeconvolve*, which, in addition to stimulus-related regressors, included regressors for linear and polynomial trends and 6 motion correction parameters. GLM parameters for the localizer blocks have been described previously (Goddard et al., 2019).

For experimental data, our model included estimates for each of the unique stimuli, and for the grey and black reference stimuli (each using 'BLOCK(12,1)'), with the 3 s fixation periods in between each stimulus used as the implicit baseline. We obtained fits for two versions of these general linear models: one in which the model fit a single beta weight for each stimulus type, and the other where it returned individual beta weights for each block, resulting in 12 beta estimates for each stimulus. The first version was used to estimate the average response of each ROI to each stimulus, and the second version was used in the classification analyses described below.

2.10. fMRI analysis: regions of interest (ROI)

We visualised results from the phase-encoded retinotopic mapping scans and the VO and hMT + functional localisers on inflated cortical representations using SUMA and used these data to define the ROIs (V1, V2, V3, V3A, V3B, LO1, LO2, hV4, VO1, VO2 and hMT+) for each participant. Across participants, we were not always able to separate V3A from V3B, or LO1 from LO2, using retinotopic mapping data and so report data for a single, combined ROI in each case (V3A/B and LO). For each of these ROIs we created a corresponding mask in the volume space, including all voxels between the grey/white boundary and the pial surface within the defined section of cortex. Finally, we also defined a ROI based on an anatomical definition of the LGN, as described by Zhang et al. (2015).

2.11. Classification analyses

We used classification analyses to measure the extent to which the pattern of BOLD responses across each ROI could be used to predict the stimulus type (excluding reference stimuli). For each classification we trained classifiers (linear support vector machine, SVM) to discriminate between two categories of block and tested on held-out data, using unnormalized beta values from the GLM as input to the classifier.

First, we trained classifiers to discriminate stimulus type and size: for 3 sizes there were 3 pairwise comparisons, and 4 stimulus types there were 6 pairwise comparisons. For these classifications we created 'pseudo-blocks' by averaging across blocks with the same value on the dimension-of-interest, but with differing values along the other dimension. Pseudo-blocks were always balanced across the irrelevant dimension-of-interest. When classifying stimulus size, each pseudo-block was the average of four blocks of the same size: one block of each stimulus type. Similarly, pseudo-blocks for classifying stimulus type were balanced across size. To ensure that our results did not depend on a particular assignment on blocks to pseudo-blocks, for every pairwise classification we generated 100 sets of pseudo-blocks, updating the random assignment of blocks for each set, and averaged classification performance across these. Second, we trained classifiers to discriminate size (3 pairwise discriminations) within each stimulus type (4 categories), and to discriminate stimulus type (6 pairwise discriminations) within each size (3 categories). For these comparisons we used unaveraged block data.

For every classification, the data comprised 24 blocks or pseudo-

blocks (12 blocks/pseudo-blocks from each of 2 categories). In each case we repeated the classification analysis 12 times, always leaving out a pair of pseudo-blocks (one from each category) and testing the accuracy of the classification rule on the held-out data. For all analyses we expressed average classifier accuracy in *d'* (a unit-free measure of sensitivity). Chance classification performance yields $d' = 0$.

2.12. Representational similarity analysis

In addition to reporting classification accuracy for decoding stimulus type and size, we also represented the pattern of classifier performance across pairs of stimulus types/sizes into dissimilarity matrices (DSMs) for each ROI. By correlating the observed DSMs with a range of model DSMs, we tested the extent to which that model could account for the observed DSM. This general approach of 'Representational Similarity Analysis' (RSA) was first applied to compare representations of object information across brain areas, and across species (Kriegeskorte et al., 2008). For decoding of stimulus type, each DSM was a 4x4 matrix, where each cell in the DSM was defined as the classification accuracy for a single pair of stimulus types. Similarly, for decoding of size, each DSM was a 3x3 matrix. The diagonal axis of these matrices was nominally zeros, and the matrix is by definition symmetric about the diagonal axis, so for all correlation values calculated below we included only the triangular part of the matrix above the diagonal.

For both stimulus type and size we tested a series of model DSMs predicting which pairs of stimuli should be the most discriminable, as illustrated below in Fig. 4A (for RSA of stimulus type) and inset in Fig. 8 (for RSA of stimulus size). Each model DSM divided the stimuli into two or three groups and predicted that classifier performance would be higher when classifying stimuli across the group boundary rather than within a group. For stimulus size, we tested three models: each where one size was in the first group, and the remaining two were in the second group. For stimulus type, we tested seven models; the first four models were based on a single stimulus type in the first group and the remaining three in the second group. The fifth model was based on stimulus chromaticity: the two chromatic stimuli (BY and RG) were in one group and the two achromatic stimuli in another. The final two models were based on stimulus preferences of the parvocellular, magnocellular and koniocellular pathways. The P/M/K model divided the 4 stimulus types into 3 groups according to these pathways: Parvocellular (RG and Ach (P-type)), Magnocellular (Ach (M-type)) and Koniocellular (BY). In the final model, the two parvocellular stimuli (RG and Ach (P-type)) are in one group and the two non-parvocellular stimuli (BY and Ach (M-type)) in another. For each participant, we correlated the observed DSMs for each ROI with each of the model DSMs, using a rank correlation (Spearman's rho), before averaging these correlation values across participants.

To visualize differences across ROIs in their coding of stimulus type, we use metric multi-dimensional scaling (MDS) to illustrate the similarities/differences across ROIs in their DSMs. For this analysis, we used the stimulus type DSMs based on data averaged across stimulus sizes and created a second-order 10x10 DSM of dissimilarity between each pair of ROIs. In the second-order DSM, each cell of the matrix was defined as $1 - \rho$, where ρ was Pearson's linear correlation between the upper diagonal portion of the stimulus type DSMs of the relevant ROIs. For the resultant second-order DSM, we used the MATLAB function *mdscale,* with criterion *'metricstress'* to identify the 2-dimensional solution with minimal stress, where stress is normalized with the sum of squares of the dissimilarities.

2.13. Statistical analyses

We performed all statistical analyses using R software (conducted in RStudio v1.1.383), and full details of the analyses are included on our OSF project (see section 'Data Availability' below). For post-hoc contrasts following repeated-measures ANOVAs, we used Cohen's $d_z = \frac{t}{\sqrt{n}}$ (see Lakens (2013), equation 7) as a measure of effect size.

3. Results

3.1. Psychophysical detection thresholds

Contrast sensitivity for each stimulus type at the smallest size is shown in Fig. 2A. The effects of stimulus size, with each subject's data normalized by their average detection threshold for the smallest stimulus are shown in Fig. 2B. Results show a consistent increase in contrast sensitivity with stimulus size for the Ach (M-type) and BY stimuli, whereas for the RG and the Ach (P-type) stimuli the effects of area summation level off for the middle and large stimulus areas. A repeated measures ANOVA of the normalized contrast sensitivities revealed significant main effects of stimulus type ($F_{(3,21)} = 23.3$, $p < 0.001$, $\omega_p^2 = 0.639$) and size ($F_{(1.14,7.95)}$[1] $= 128.1$, $p < 0.001$, $\omega_p^2 = 0.855$), and a significant interaction ($F_{(6,42)} = 30.7$, $p < 0.001$, $\omega_p^2 = 0.650$) between these effects. Pairwise comparisons (with Tukey's HSD correction for multiple comparisons) reveal that for the largest stimulus size there is no significant difference between normalized contrast sensitivity for the RG and Ach (P-type) stimuli ($t_{(52)} = 0.69$, $p = 0.901$, Cohen's $d_z = 0.24$), and the small difference between detection thresholds for the Ach (M-type) and BY stimuli failed to reach statistical significance ($t_{(52)} = 2.57$, $p = 0.061$, Cohen's $d_z = 0.91$), but there are significant differences for the remaining four pairwise comparisons ($t_{(52)} > 8.69$, $p < 0.001$, Cohen's $d = 3.07$–4.23, in each case).

This pattern of results is consistent with the RG and Ach (P-type) stimuli favouring neural mechanisms that are relatively more concentrated near the fovea, with a steeper drop off with eccentricity than for the mechanisms underlying detection of the Ach (M-type) or BY stimuli, consistent with previous literature (Allen and Hess, 1992; Anderson et al., 1991; Mullen and Kingdom, 2002).

3.2. Classification analyses

We measured the fMRI (BOLD) signals induced by the four stimulus types, at each of the three sizes, at suprathreshold contrasts. We used classification analyses to test for evidence that the pattern of response across voxels in each ROI varied reliably with stimulus type and size. We trained and tested classifiers on their ability to discriminate either stimulus size or stimulus type using the pattern of responses across voxels within each ROI. Classifier performance (Fig. 3) was above chance for decoding stimulus type and size in each ROI ($t_{(7)} \geq 3.92$ in each case, $p < 0.01$, FDR corrected for multiple comparisons) with the exception of the decoding of size in LGN ($t_{(7)} = -1.03$, $p = 0.831$). A 2-way repeated measures ANOVA of the effect of ROI and classified feature (stimulus type vs size) on classifier accuracy revealed significant main effects of both ROI ($F_{(9,63)} = 41.50$, $p < 0.001$, $\omega_p^2 = 0.905$) and classified feature ($F_{(1,7)} = 46.88$, $p < 0.001$, $\omega_p^2 = 0.577$), as well as a significant interaction between these effects ($F_{(9,63)} = 41.76$, $p < 0.001$, $\omega_p^2 = 0.696$).

Post-hoc pairwise contrasts (with Tukey's HSD correction for multiple comparisons) revealed that for most ROIs there was a significant ($p < 0.01$) difference in classifier accuracy between stimulus size and type, although the direction of this effect varied across ROIs. For earlier ROIs, decoding of size tended to be more accurate than decoding of stimulus type, including for V1 ($t_{(42)} = 11.27$, Cohen's $d_z = 3.98$), V2 ($t_{(42)} = 10.44$, Cohen's $d_z = 3.69$), V3 ($t_{(42)} = 8.97$, Cohen's $d_z = 3.17$), V3A/B ($t_{(42)} = 5.51$, Cohen's $d_z = 1.95$), LO ($t_{(42)} = 6.10$, Cohen's $d_z = 2.16$) and hV4 ($t_{(42)} = 5.07$, Cohen's $d_z = 1.79$). Conversely, for areas hMT ($t_{(42)} = 3.45$, Cohen's $d_z = 1.22$) and VO2 ($t_{(42)} = 3.21$, Cohen's $d_z = 1.13$), accuracy for decoding stimulus type was higher than for decoding size. For area VO1, there was no significant difference between decoding of stimulus type and size.

Overall, the comparison of classifier accuracy between size and

[1] Degrees of freedom corrected using a Greenhouse-Geisser correction for violations of sphericity.

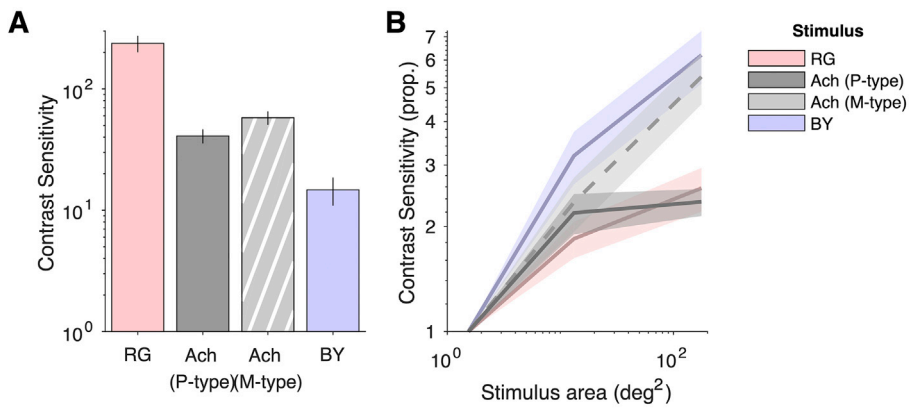

Fig. 2. Detection thresholds across stimulus size. In **A**, contrast sensitivity (1/detection threshold, with detection threshold in cone contrast units) is shown for the smallest stimulus area (1.5 deg2) for each stimulus type, with contrast sensitivity on a logarithmic scale. In **B**, contrast sensitivity is shown across the three stimulus sizes, normalized by the sensitivity for the 1.5-deg2 stimulus, on log-log axes. In both plots, error bars indicate the 95% confidence intervals of the between-subjects mean ($n = 8$).

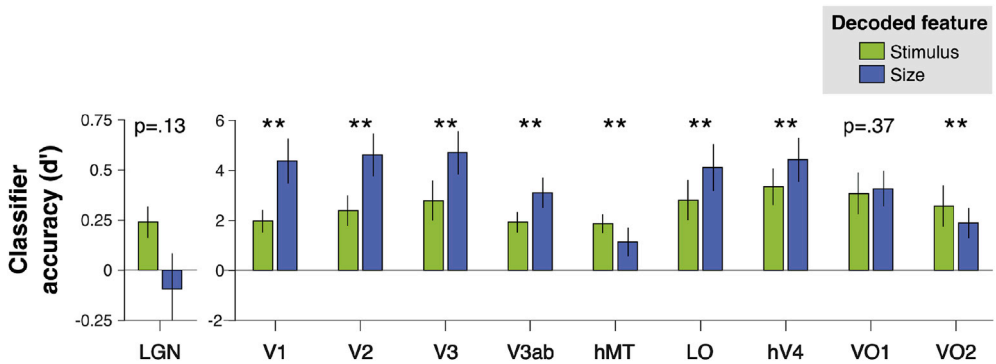

Fig. 3. Classification analysis: stimulus type and size. We used a classification analysis to ask how well the pattern of BOLD responses (beta weights) within each ROI could be used to predict the size and type of the stimulus. Error bars indicate the 95% confidence intervals of the between-subjects mean ($n = 8$). Asterisks show where the ROI has significantly different classifier accuracy when decoding stimulus vs size ($p < 0.01$).

stimulus type provides a relative metric for revealing trends across areas as, typically, classifier accuracy is not directly comparable between ROIs. Here the relative transition from better decoding of size to better decoding of stimulus type likely reflects the increase in receptive field size from earlier to later visual cortex (Dumoulin and Wandell, 2008; Harvey and Dumoulin, 2011), combined with a magnification of the central part of the visual field that continues to increase throughout the visual processing hierarchy. For example, Schira et al. (2009) found that the area of the cortical surface responding to the central 0.75 degs of visual field was larger in areas V2 and V3 than in area V1, despite these ROIs being smaller in overall area. Harvey and Dumoulin (2011) report greater foveal representation in ventral areas hV4 and LO compared to earlier areas (V1–V3). These factors will reduce the impact of changes in the retinotopic extent of the stimuli.

Interestingly, for LGN, although classifier accuracy was lower overall, classifier performance was above chance for decoding of stimulus type and tended to be higher than for decoding of stimulus size (which was at chance performance), although the difference between decoding of stimulus type vs size did not reach statistical significance ($t_{(42)} = 1.56$, $p = 0.13$, Cohen's $d_z = 0.55$). The failure of classifiers to decode stimulus size from LGN, despite above chance decoding of stimulus type, could be related to the observation in previous work that visual field eccentricity maps measured with fMRI tend to be poorly defined in LGN (DeSimone et al., 2015; Schneider et al., 2004). It also relates to results later in this section showing a lack of area summation in the LGN.

3.3. *Differences in stimulus coding across area: representational similarity analysis*

The classification of stimulus type in Fig. 3 is based on classifier performance averaged across all stimuli, whereas in fact some stimulus pairs may be better discriminated than others. Moreover, classifier performance cannot be directly compared across visual areas as it is influenced by many different factors. To better characterize differences across ROIs in how they respond to our stimuli, and to determine which stimulus pairs drive the classifier better or worse, we performed a representational similarity analysis (RSA) (Kriegeskorte et al., 2008) on the pattern of pairwise classifications across stimulus types. For each participant's data, we rank correlated each ROI's pattern of classification performance with each of 7 models, shown in Fig. 4. Each of these models predicts the pairs of stimuli that will be more discriminable (yellow cells Fig. 4A) and less discriminable (blue cells in Fig. 4A). By comparing how well these models account for the observed patterns of results, we can identify transformations in how visual information is represented across different ROIs.

The four models in upper row of Fig. 4A are each based on the predicted pattern of classification performance if there were a single stimulus type that is different to the other three, and this difference is dominating the observed pattern of results. As shown in Fig. 4C, these four models are each slightly negatively correlated with each other, meaning that positive correlation with one model will tend to produce negative correlation with the remaining models.

The remaining three models (lower row in Fig. 4A) are each based on different groupings of the stimuli. The chromatic/achromatic ('Chr/Ach') model predicts that the presence or absence of chromatic or achromatic contrast in a stimulus pair is dominant in driving classifier performance. Under this model, the chromatic stimuli (RG and BY) are predicted to be poorly discriminated from one another, as are the two achromatic stimuli, but any pairing of a chromatic with an achromatic stimulus is predicted to produce good discrimination. Similarly, the 'P/

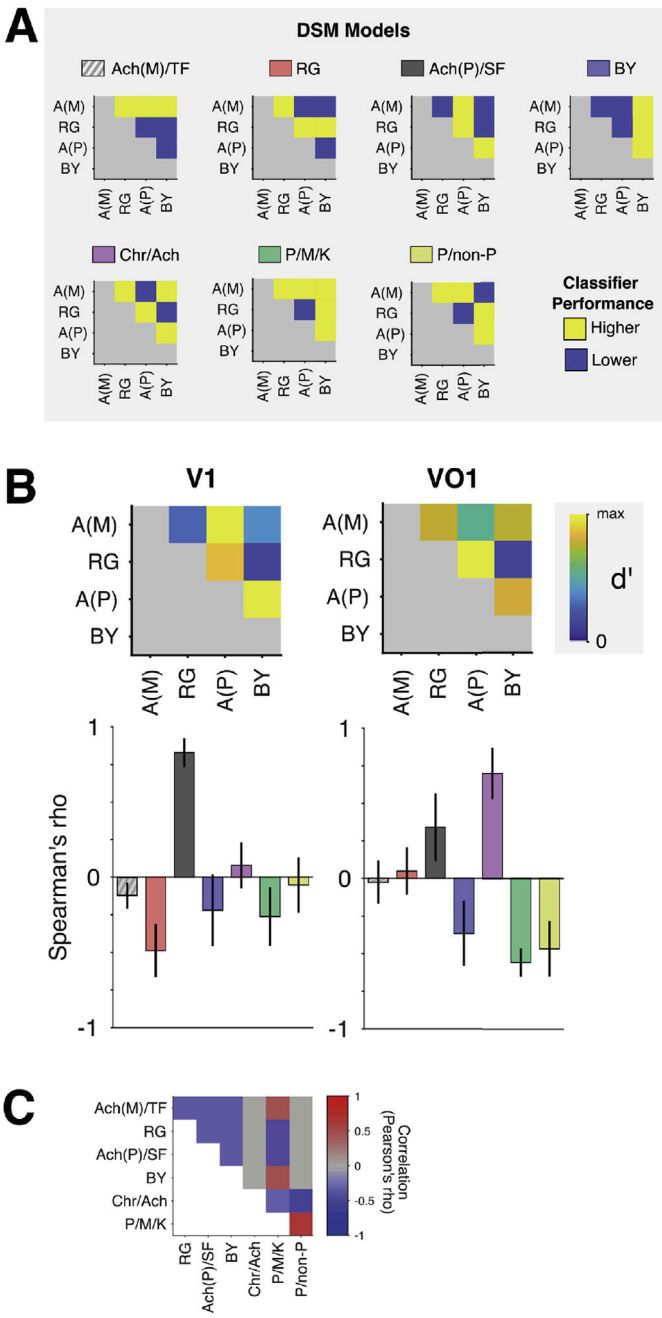

Fig. 4. Representational Similarity Analysis (RSA) for stimulus type. The pattern of classification performance across stimulus pairs (shown in each dissimilarity matrix, or DSM) was rank correlated with each of 7 models (**A**). The first four models (upper row in **A**) predict the pattern of accuracies when a single stimulus type evokes a response that is dissimilar to all others. The remaining three models (lower row in **A**) are: 'Chr/Ach' (chromatic/achromatic) that predicts best performance when one stimulus of the pair is chromatic and the other achromatic; 'P/M/K', that predicts best performance when each stimulus of a pair is thought to be mediated by a different LGN layer: Parvocellular (RG and Ach (P)), Magnocellular (Ach (M)) and Koniocellular (BY); 'P/non-P' that predicts best performance when one member of the pair is a parvocellular-based stimulus (RG or Ach (P)) and the other a non-parvocellular stimulus (BY and Ach (M)). In **B**, example data are shown for two ROIs (V1 and VO1), including the average DSM and the average correlation of the DSM with each model. Error bars indicate the 95% confidence intervals of the between-subjects mean ($n = 8$). Correlations between models (**C**) show that the first four models are weakly negatively correlated with one another, but orthogonal to the Chr/Ach and P/non-P models.

M/K' and 'P/non-P' models predict good between-group discrimination and poor within-group discrimination when the stimuli are grouped either according to their parvocellular, magnocellular or koniocellular neural bias (P/M/K model), or their parvocellular versus non-parvocellular bias (P/non-P model). Fig. 4C shows that the P/M/K and P/non-P models are highly correlated with each other, but both negatively correlated with the Chr/Ach model.

Fig. 4B illustrates the result of this analysis for two example ROIs (V1 and VO1). Area V1 shows highest correlation with the model based on the Ach (P-type) stimulus, which is seen both in the model correlation values as well as in the visual similarity between the Ach(P)/SF model and the actual average dissimilarity matrix found for V1. This suggests that for area V1, the Ach (P-type) stimulus produces a voxel response pattern that is most distinctive, while the remaining three stimuli produce more similar responses patterns. Since the Ach (P-type) stimulus differs from the other three in its spatial frequency (5 versus 0.5 cycles/deg), this model is equivalent to a prediction based on responses to spatial frequency. The Ach (M-type) model is similarly equivalent to a prediction based on temporal frequency. For the second example in Fig. 4B, area VO1, the highest correlation is for the Chr/Ach model, showing that the average dissimilarity matrix found for VO1 is best predicted by the presence of chromatic vs achromatic contrast in the stimuli. There is still a positive correlation with the Ach(P)/SF model but this is weaker than for the Chr/Ach model and weaker than the correlation found in V1, suggesting a shift from a voxel response pattern that is more distinctive for the Ach (P-type) contrast in V1 to one reflecting chromatic vs achromatic contrast in VO1.

In Supplementary Fig. 1, model correlations are shown for all models across all ROIs. Four of the seven models we tested did not show above chance positive correlation with any ROI: these were the models based on the RG or BY being most dissimilar, and the P/M/K and P/non-P models. Neither the P/non-P nor the P/M/K models provided a good account of the data in the LGN, despite being based on the predicted responses of the LGN layers. The three remaining models were the Ach(M)/TF, Ach(P)/SF and Chr/Ach models, which each performed well in different ROIs. Of these three models, the Ach(M)/TF and Ach(P)/SF models are weakly negatively correlated with each other, while the Chr/Ach model is independent of the other two (see Fig. 4C). To differentiate between the two achromatic models, we restricted our remaining analyses to these three top-performing models, and for the two achromatic models we used partial correlations to measure the model's ability to account for the observed dissimilarity matrices. Specifically, we measured the performance of the Ach(M)/TF model as the partial correlation between the data and this model (controlling for correlation with the Ach(P)/SF model), and similarly we measured the performance of the Ach(P)/SF model as the partial correlation between the data and this model (controlling for correlation with the Ach(M)/TF model). The results of this analysis are plotted in Fig. 5.

To test for differences in how well these 3 models accounted for the data across different ROIs we performed a 2-way repeated measures ANOVA to test for the effect of model type and ROI on correlation/partial correlation value (Spearman's rho). This revealed a significant main effect of ROI ($F_{(9,63)} = 3.86, p < 0.001, \omega_p^2 = 0.022$), but no significant main effect of model type ($F_{(2,14)} = 2.85, p = 0.092, \omega_p^2 = 0.056$), and a significant interaction between these effects ($F_{(18,126)} = 5.61, p < 0.001, \omega_p^2 = 0.257$). To identify which models performed significantly better than other models for each ROI, we performed pairwise comparisons (with Tukey HSD correction for multiple comparisons). Below we summarize the main findings of these pairwise comparisons. Full results of these pairwise comparisons, including effect sizes, are given in Supplementary Table 1.

In the LGN (Fig. 5A), classifier performance was lower and model correlation values showed more inter-subject variability than for the cortical areas. The Ach(M)/TF and Ach(P)/SF models had higher performance than the Chr/Ach model, but no pairwise comparisons for differences between models reached significance. The positive

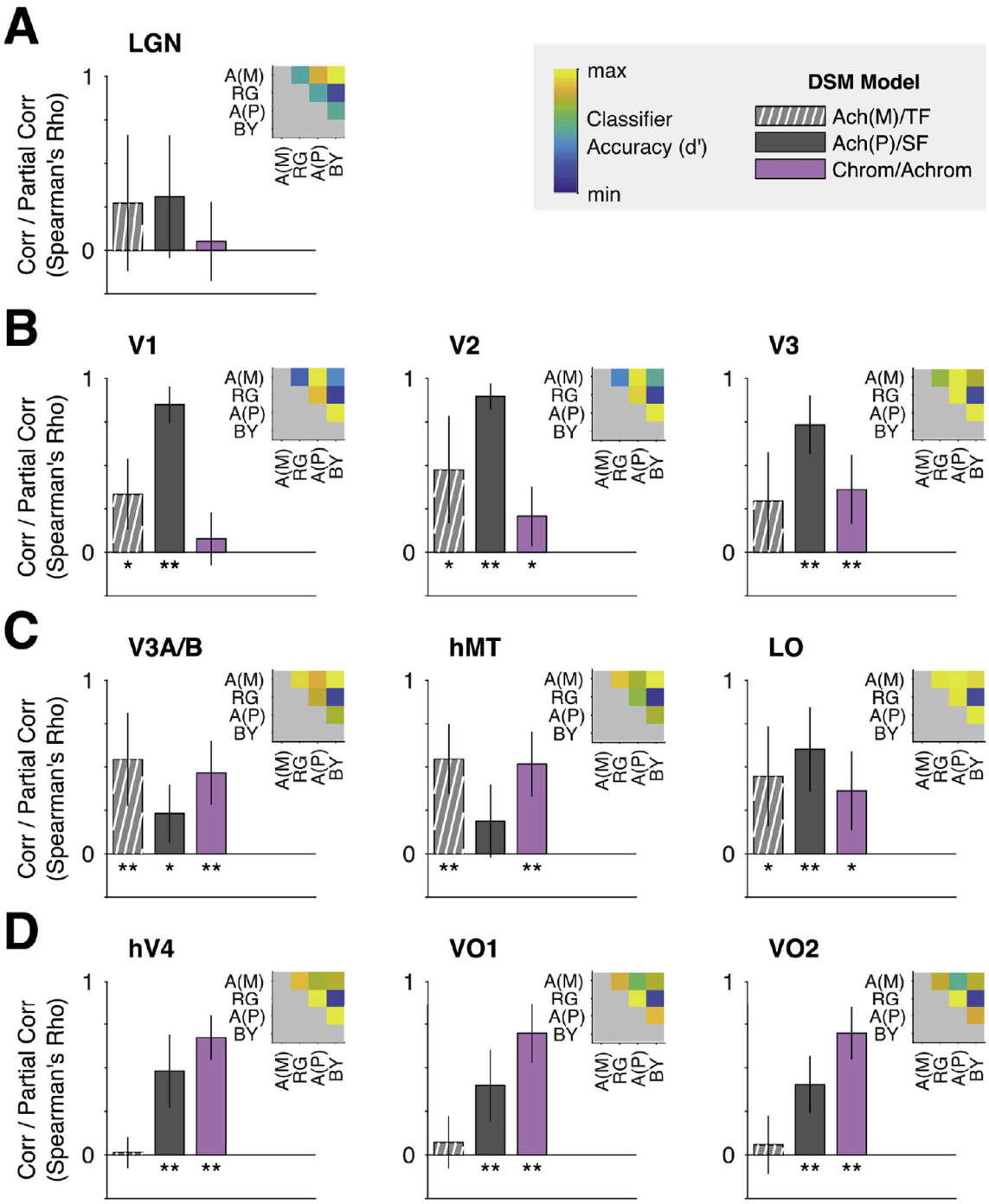

Fig. 5. Response similarity analysis (RSA): stimulus type, for the LGN (A), and early (B), dorsal/lateral (C) and ventral (D) visual areas. For each region of interest, the bar plots show the average correlation of each participant's DSM with each of the three top-performing models, and the inset shows the average DSM (described in Fig. 4). Unlike in Fig. 4, here correlations with the two achromatic models are partial correlations: the partial correlation with the Ach(M)/TF model (controlling for correlation with the A(P)/SF model) and partial correlation with the Ach(P)/SF model (controlling for correlation with the A(M)/TF model). Correlations with the Chr/Ach model are also independent from the two achromatic models, since the models are independent (see Figure 4C). All error bars indicate the 95% confidence intervals of the between-subjects mean ($n = 8$), and asterisks indicate where the correlation is significantly above 0 (**$p < 0.01$, *$p < 0.05$, FDR corrected for multiple comparisons).

correlation with the Ach(M)/TF model suggests that the observed pattern of classifier performance could reflect a grouping of magnocellular vs non-magnocellular responses, although as noted above, the models based on other layer-based parcellations of the LGN performed poorly.

In the early visual areas (V1–V3, Fig. 5B), the Ach(P)/SF model significantly outperformed the other two models ($p < 0.05$ in each case), except in V3, where it did not significantly outperform the Chr/Ach model. In each of these early visual areas, there was no statistically significant difference between the Ach(M)/TF and Chr/Ach models (see Supplementary Table 1).

In the dorsal visual areas V3A/B and hMT (Fig. 5C), the Ach(M)/TF model and the Chr/Ach model are the two top-performing models, but there were no statistically significant pairwise comparisons between models (see Supplementary Table 1). The Ach(P)/SF model did not correlate particularly well with the data, unlike in the early visual areas. This suggests that despite the robust BOLD responses of V3A/B and hMT

to the high spatial frequency achromatic stimulus (see below), these areas carry relatively more information about the temporal frequency and the achromatic vs. chromatic contrast of the stimuli than they do about the spatial form. It is also notable that for hMT there is a dissociation between the stimuli that evoked the largest average responses (see below), and the stimuli that the RSA suggests are most distinctive in terms of the pattern of responses across voxels that they induce. The high performance of the Ach(M)/TF model in areas V3A/B and hMT is consistent with these areas carrying information about stimulus temporal frequency and motion (Singh et al., 2000; Tootell et al., 1997).

In area LO, the three models performed similarly well, and no pairwise comparisons were significant. In the ventral visual areas (hV4, VO1 and VO2, Fig. 5D), the Chr/Ach model was the best performing model, significantly outperforming the Ach(M)/TF model in each area. In each case, the second-best performing model was the Ach(P)/SF model, which significantly outperformed the Ach(M)/TF in area hV4 only (see Supplementary Table 1).

A trend analysis across early to late ROIs along the ventral visual pathway (areas V1, V2, V3, hV4, VO1 and VO2) revealed significant linear trends (with Bonferroni correction for multiple comparisons) in the correlation with each of the three models (Ach(M)/TF model: $t_{(84)} = -4.05, p < 0.001$, Cohen's $d_z = -1.43$; Ach(P)/SF model: $t_{(84)} = -5.59, p < 0.001$, Cohen's $d_z = -1.98$; Chr/Ach model: $t_{(84)} = 6.89, p < 0.001$, Cohen's $d_z = 2.44$). That is, moving along the ventral visual stream the correlation with the Ach(P)/SF and Ach(M)/TF models decreased while the correlation with the Chr/Ach model increased. Note that with the partial correlations each of these three models are orthogonal to one another, so an increase in correlation with one model does not predict a decrease (or increase) in correlation with any other. These three trends suggest that when moving from early to later areas along the ventral visual pathway there is relatively less information about stimulus spatial and temporal frequency, but more information about whether the stimulus contains chromatic or achromatic contrast.

Differences between ROIs in their coding of stimulus type are also illustrated by the second-order MDS solution in Fig. 6. This re-emphasizes the trends discussed above: there is a progression in stimulus coding along the visual pathway from area V1 to later areas, and a separation between the stimulus coding of ventral and dorsal visual areas, with the ventral areas hV4, VO1 and VO2 clustered together, and dorsal areas hMT and V3A/B clustered together.

3.4. Differences in cortical magnification of the fovea do not account for inter-area differences: RSA of stimulus type across sizes

Responses to the stimuli we use here vary with eccentricity in different ways, as demonstrated in previous work (Allen and Hess, 1992; Anderson et al., 1991; Mullen and Kingdom, 2002) and reflected in the detection thresholds measured here (Fig. 2). Fig. 2 implies that the neural mechanisms responding to the RG and Ach (P-type) stimuli are relatively more concentrated in central vision, with a steeper drop off with eccentricity than for the mechanisms underlying detection of the Ach (M-type) or BY stimuli, suggesting that increasing stimulus size may positively affect the correlations with the Ach (M-type) or BY models more than the other two. Additionally, cortical magnification of the fovea generally increases when moving from V1 to higher-level areas (e.g. Harvey and Dumoulin, 2011; Schira et al., 2009), potentially giving neural mechanisms biased for central vision greater representation in higher over lower areas. To test whether the stimulus coding biases we find in Fig. 5 are, in general, dependent on stimulus size, we performed RSAs on the decoding of stimulus type for each stimulus size. For these analyses we used classifier accuracies from the pairwise discriminations of stimulus type, performed on fMRI data from within each stimulus size.

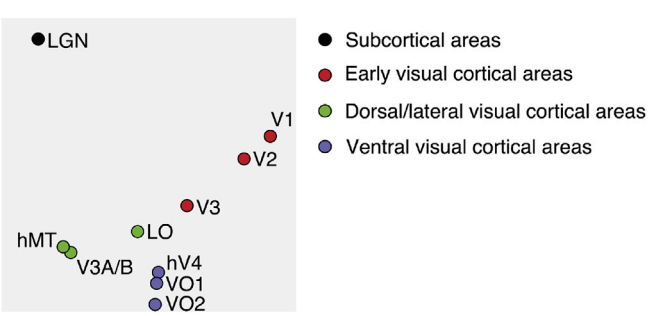

Fig. 6. Differences across ROIs in coding of stimulus type. Each region of interest is shown in a 2- dimensional space (unitless dimensions) produced by metric MDS applied to pairwise inter-ROI correlations in the stimulus-type DSMs illustrated in Fig. 5. In this plot, proximity between any pair of ROIs reflects the similarity of their encoding of stimulus type.

In Fig. 7, we plot each ROI's model correlations/partial correlations across stimulus size for the same three models as in Fig. 5. We performed a 3-way repeated measures ANOVA to test for the effect of stimulus size, model type and ROI on correlation/partial correlation value (Spearman's rho). As for the main ANOVA (see above), there was a significant main effect of ROI ($F_{(9,63)} = 4.68, p < 0.001, \omega_p^2 = 0.038$), as well as a significant interaction between ROI and model ($F_{(18,126)} = 5.01, p < 0.001, \omega_p^2 = 0.126$), but here there was also a significant main effect of model type ($F_{(2,14)} = 6.80, p = 0.009, \omega_p^2 = 0.070$). There was no significant main effect of stimulus size ($F_{(2,14)} = 0.25, p = 0.786, \omega_p^2 = 0.001$), but there was a significant 3-way interaction between stimulus size, model type and ROI ($F_{(36,252)} = 2.35, p < 0.001, \omega_p^2 = 0.063$). To identify which changes in correlation with stimulus size were driving this 3-way interaction, we tested for a linear trend with stimulus size for each of the models plotted in Fig. 7, for each ROI. No linear trends with stimulus size survived a Bonferroni correction for multiple comparisons, but a number of ROIs showed trends which approached significance ($p < 0.05$, uncorrected).

For correlation with the Ach(M)/TF model, areas hMT and VO2 each showed some evidence of linear trends across stimulus size ($t_{(327)} = 2.64, -2.66$; $p = 0.009, 0.008$ (uncorrected); Cohen's $d_z = 0.93, -0.94$ respectively). In VO2 the trend was for decreasing correlation with increasing stimulus size, inconsistent with our prediction, and at each size correlation with this model was low, making this trend harder to interpret. In hMT this trend was for increased correlation with increasing stimulus size, consistent with our prediction. This hMT trend could imply that areas with high foveal magnification (e.g. ventral areas) may have been less likely to show a positive correlation with the Ach(M)/TF model. However, there are two counter-examples which we think make it unlikely that foveal magnification masked an Ach(M)/TF-type effect in these areas. First, areas V3A/B and hMT also have relatively high foveal magnification, yet these areas showed high correlation with the Ach(M)/TF model. Second, we found relatively low correlation with the Ach(M)/TF model in area V1, with the lowest foveal magnification of these areas.

For the Ach(P)/SF model, based on the particularly steep drop-off in parvocellular mechanisms with eccentricity and the shift in SF tuning towards lower frequencies with eccentricity for areas V1 to hV4 (Henriksson et al., 2008), we predicted that correlation would tend to be highest for the smallest stimulus size. As stimulus size increases, the overall signal available to the classifier will increase, but the extra stimulus area will recruit parts of the visual field that are less sensitive to high spatial frequencies. There were some non-significant effects in this direction for areas hV4, VO1 and VO2, including an effect in VO2 which approached significance ($t_{(327)} = -1.96, p = 0.051$ (uncorrected), Cohen's $d_z = -0.69$). However, the only area with a linear trend in correlation with the Ach(P)/SF model that reached significance was area V2 ($t_{(327)} = 2.14, p = 0.033$ (uncorrected), Cohen's $d_z = 0.76$), and this trend was in the opposite direction to our prediction: the correlation with the Ach(P)/SF model increased with stimulus size. This does not rule out the possibility that correlation with the Ach(P)/SF model is affected by receptive field size, but it suggests that the receptive field size is not a strong predictor of correlation with this model.

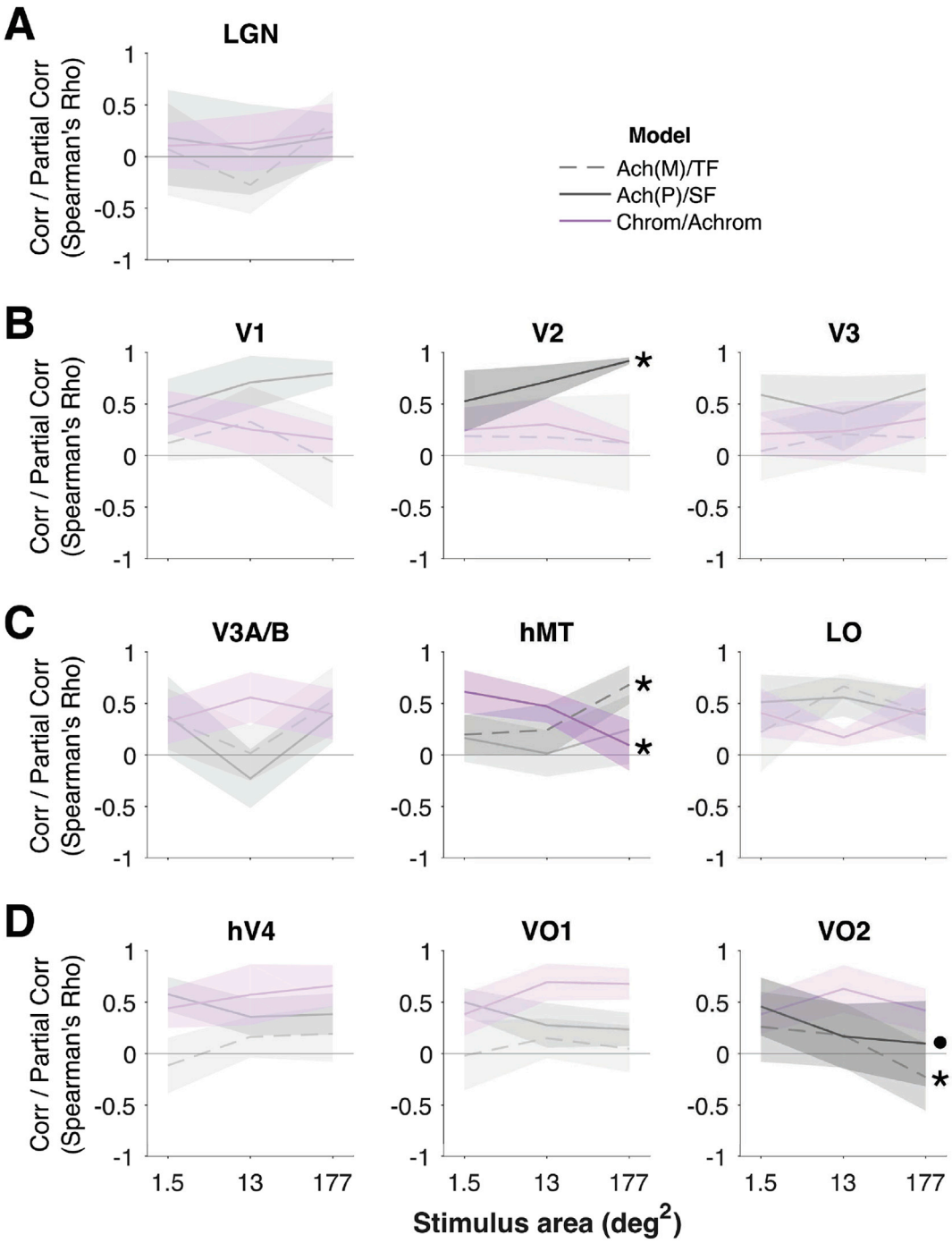

Fig. 7. Response similarity analysis (RSA): stimulus type, within each stimulus size, for the LGN (**A**), and early (**B**), dorsal/lateral (**C**) and ventral (**D**) visual areas. Lineplots show the average correlation/partial correlation of each participant's DSM for the same models is in Fig. 5. Cases where there was a linear trend in model correlation across stimulus size are highlighted with asterisks ($p < 0.05$, uncorrected) or filled circles ($p = 0.051$, uncorrected), remaining cases are plotted with reduced contrast. Shaded error bars indicate the 95% confidence intervals of the between-subjects mean ($n = 8$).

We did not have a clear *a priori* prediction for how correlation with the Chr/Ach model would vary with stimulus size. Across ROIs, correlation with the Chr/Ach model increases (Fig. 5), however, there is no significant effect of stimulus size in most areas. The only area with a significant linear trend in correlation with this model across stimulus size was area hMT ($t_{(327)} = -2.83$, $p = 0.005$ (uncorrected), Cohen's $d_z = -1.00$), for which model correlation decreased with increasing stimulus area. Across remaining areas there was inconsistency in the (non-significant) trends, with no discernable relationship between cortical magnification factor and model correlation. For these reasons, we think it unlikely that the Chr/Ach model is directly related to increases in receptive field size or cortical magnification factor across ROI. Instead, the increasing correlation with the Chr/Ach model from early to later visual areas reported above suggests an increasing distinctiveness of neural responses to chromatic and achromatic stimuli.

In summary, the results of our RSA analyses (Fig. 5) are quite robust across stimulus size, with (uncorrected) significant effects on the best performing models found only in V2, hMT, and VO2. It is possible that these and some non-significant linear trends with stimulus size are genuine effects that our sample size ($n = 8$) was too small to detect.

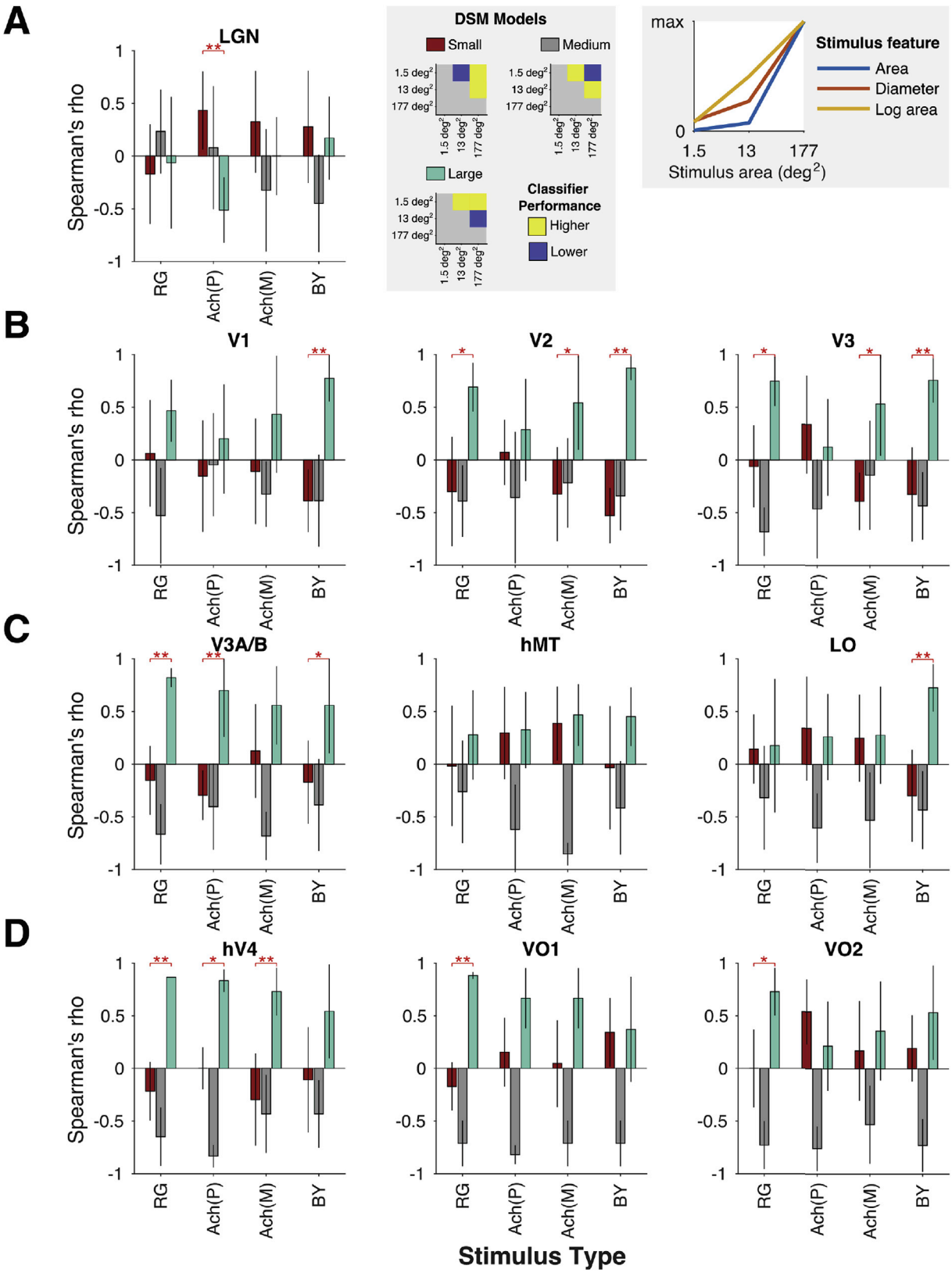

Fig. 8. Response similarity analysis (RSA): stimulus size. Plotting conventions as in Fig. 5. Left inset (top middle) illustrates the three alternative models for pairwise decoding of stimulus size, where each predicts a single size is most dissimilar to the other two. Right inset (top right) shows how stimulus diameter, area and log area varied across the 3 stimulus sizes. Red asterisks indicate cases where there is a significant difference between the 'Small' and 'Large' models in their correlation with the data (*$p < 0.05$; **$p < 0.01$). In terms of diameter and area, the largest stimulus size is the most dissimilar, which predicts highest correlation with the 'Large' model. In terms of log area, the stimulus sizes are equally spaced, predicting similar correlation with the 'Small' and 'Large' models. Higher correlation with the 'Small' model than with the 'Large' model suggests a magnification of the foveal visual field that is greater than log scaling.

Despite this, the inter-area differences we found in stimulus representation and the directions of these sometimes-weak trends across stimulus size suggest trends from early to ventral visual areas cannot be accounted for by differences between areas in their receptive field size or their magnification of the fovea.

3.5. Inter-stimulus differences in size coding: RSA of stimulus size

Another way that differing responses with eccentricity could be reflected in our data is in the decoding of stimulus size. To test this, we also performed RSA on the decoding of size for each ROI (Fig. 8). For this analysis we used classifier accuracies from the pairwise discriminations of stimulus size, performed on fMRI data from within each stimulus type. In this case there were only 3 models, where each model was based on stimuli of a single size evoking the most distinctive responses (illustrated in the inset in the top middle of Fig. 8). High correlation with the 'Small' model indicates that the ROI's response pattern shows greater change when the stimulus area increases from 1.5 to 13 degs^2 than it does for 13 degs to 177 degs^2, while high correlation with the 'Large' model indicates the reverse. If, for example, the ROI response patterns were linearly related to stimulus diameter or area, then the response to the largest stimulus would be most dissimilar (see inset in top right of Fig. 8). If, however, ROI response were logarithmically related to stimulus area (yellow line in inset of Fig. 8), then the stimulus sizes used here are approximately equally spaced, which predicts that the Small and Large models will be correlated with the data to the same extent. Where correlation with the Small model exceeds that of the Large model this suggests a magnification of the foveal visual field that is greater than log scaling. If responses are proportional to the psychophysical detection for these stimuli found in Fig. 2B, then for the BY and Ach (M-type) stimuli there should be approximately equal correlation with the Small and Large models, while for the RG and Ach (P-type) stimuli the Small model should outperform the Large model.

Note that the 'Medium' model is based on the unrealistic prediction that the mid-sized stimulus is the most distinctive: if the ROI's response is monotonically varying with size (as predicted in all the scenarios listed above) then the classification of 1.5 vs 13 deg^2 stimuli should produce the highest classifier performance, leading to negative correlation with the Medium model. In this way, the extent to which the data are negatively correlated with the Medium model gives an indication of how consistently the ROI's data were varied in an orderly way with stimulus size.

A 3-way repeated measures ANOVA on the correlation values with the 'Small' and 'Large' models (excluding the Medium model), revealed a significant main effect of model type ($F_{(1,2)} = 218.9$, $p = 0.005$, $\omega_p^2 = 0.238$), and a significant interaction between ROI and model type ($F_{(9,18)} = 4.26$, $p = 0.004$, $\omega_p^2 = 0.066$), but no significant main effects of ROI ($F_{(9,18)} = 1.53$, $p = 0.213$, $\omega_p^2 = 0.018$) or stimulus type ($F_{(3,6)} = 0.12$, $p < 0.943$, $\omega_p^2 = 0.010$). For each stimulus type, we tested for a significant difference between the Small and Large models (averaged across ROIs). Contrary to the prediction from psychophysical thresholds, there was no stimulus where the Small model consistently had higher correlation than the Large model. Instead, in each case, the Large model outperformed the Small model, with the greatest average difference for the BY stimulus (diff = 0.67, $t_{(27)} = 5.30$, $p < 0.001$, Cohen's $d_z = 1.87$), then the RG stimulus (diff = 0.64, $t_{(27)} = 4.99$, $p < 0.001$, Cohen's $d_z = 1.76$), the Ach (M-type) stimulus (diff = 0.43, $t_{(27)} = 3.35$, $p = 0.002$, Cohen's $d_z = 1.18$) and the smallest difference for the Ach (P-type) stimulus (diff = 0.13, $t_{(27)} = 1.00$, $p = 0.327$, Cohen's $d_z = 0.35$). Across stimuli, this ordering of differences between the Large and Small models is partly in line with behavioral sensitivity: in line with predictions, the BY stimulus has the strongest preference for the Large model, and the Ach(P)/SF model has the smallest.

Pairwise comparisons between the correlation with the Small and Large models for each ROI and stimulus type (with Tukey's HSD correction for multiple comparisons) revealed the pairwise differences highlighted in red in Fig. 8. For all cortical ROIs (Fig. 8, B-D), data for most stimulus types correlated more strongly with the Large model than for the Small model. A clear exception is the LGN, where the Small model significantly outperformed the Large model for the Ach (P-type) stimulus, consistent with a strong foveal bias in the LGN.

3.6. Stimulus differences in mean BOLD signal change

As a final characterization of the responses evoked by these stimuli, we plot each ROI's average BOLD signal change evoked by each stimulus at each size (Fig. 9). Although we expect that our classifier analyses and RSAs to be more sensitive for detecting response differences, we included this measure of univariate change to see how average responses related to classifier accuracy, RSA model correlation, stimulus cone contrast, and the detection threshold of the stimuli. For this analysis, we performed a contrast comparing the responses to the black (dark) reference blocks with responses to all other block types to define a localizer for visually responsive voxels. The results in Fig. 9 are based on the average responses of all voxels for which this contrast showed a significant difference, with a liberal criterion of $t_{(768)} > 1.65$ (two-sided test, equivalent to $p < 0.10$, uncorrected).

A 3-way repeated measures ANOVA (9 ROIs x 4 stimulus types x 3 stimulus sizes) revealed significant main effects on the average beta values of ROI ($F_{(9,54)} = 14.02$, $p < 0.001$, $\omega_p^2 = 0.786$) and size ($F_{(2,12)} = 49.25$, $p < 0.001$, $\omega_p^2 = 0.732$), but not stimulus type ($F_{(3,18)} = 1.98$, $p = 0.153$, $\omega_p^2 = 0.204$). To better characterize these differences across ROIs, for each ROI, we performed a 2-way repeated measures ANOVA on the effects of stimulus type and size on the beta values. For each ROI, with the exception of the LGN, there was a significant main effect of stimulus size ($p < 0.01$). In each area response amplitude (average beta) tended to increase with stimulus size. For areas V1, V2, V3, LO, hMT and hV4 there was also a significant main effect of stimulus type ($p < 0.05$). The response differences driving these main effects varied across areas: in V1, V2, V3 and hV4, the Ach (P-type) stimulus induced the largest response amplitudes, while the RG stimulus induced the smallest. In area hMT, the RG stimulus induced an unexpectedly large response. Although it seems counter-intuitive that hMT has a stronger response to RG contrast than Ach, similarly strong responses have been reported previously for sine-wave stimuli that are temporally counterphasing (Liu and Wandell, 2005; Mullen et al., 2010). The known response of M-cells to RG chromatic modulation may contribute to the BOLD response in hMT (Lee et al., 1989).

There was no ROI for which these amplitude differences were clearly related to either cone contrast, or to multiples of detection threshold. The RG stimulus had the lowest cone contrast (4%) and the two achromatic stimuli had the highest (50%), which is the approximate ordering for early visual areas, except that the response to the Ach (P-type) stimulus was larger than to the Ach (M-type) stimulus, even though they were matched in cone contrast. The BY stimulus was the lowest contrast in terms of multiples of detection threshold (3-16 x), while the Ach (M-type) was the highest (27-146 x) yet there was no ROI for which this ordering was reflected in BOLD amplitudes.

Only V1, V2, and LO showed a significant interaction between stimulus size and type ($p < 0.05$). Across each of V1, V2 and LO, the response to the Ach (P-type) stimulus increased from 13 to 177 degs^2 to a greater extent than for the remaining three stimuli. In areas V1 and V2 there was a smaller increase in response from 13 to 177 degs^2, while for LO the average responses to the 177 degs^2 stimuli were lower than for the 13 degs^2 stimuli. Note that this interaction is in the opposite direction to that predicted by the detection thresholds: based on the psychophysical data we would predict that the Ach (M-type) and BY stimuli would show a greater increase with stimulus size. We discuss this lack of correspondence further below. Full results of these analyses, including effect sizes, are shown in Supplementary Table 2.

In this study, there have been many examples demonstrating the disassociation between the amplitude of BOLD response to stimuli and

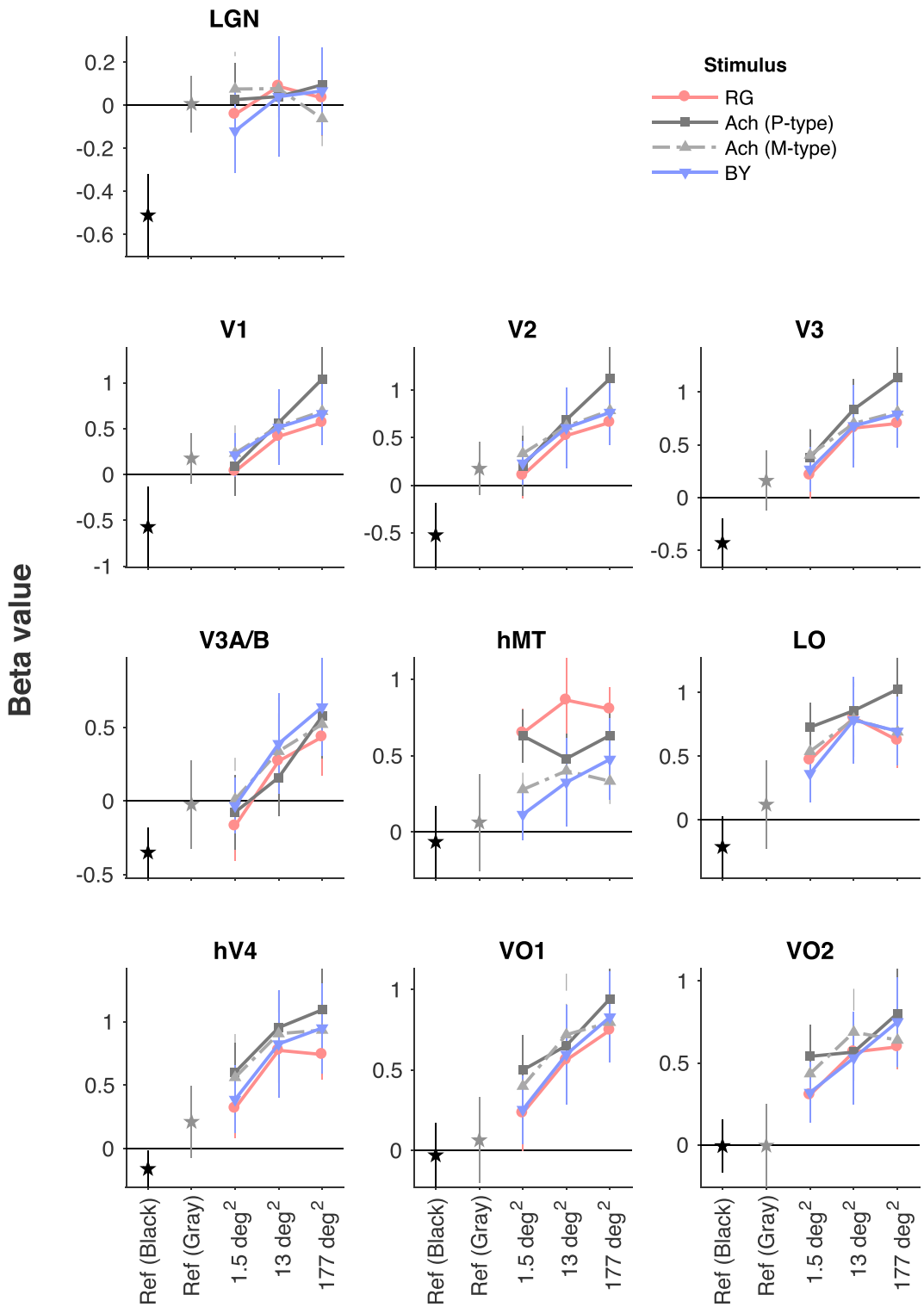

Fig. 9. BOLD response amplitudes across stimulus type and size. Beta values are the estimate of BOLD response (equivalent to % signal change) obtained from the GLM. For each ROI, beta values are average across all visually responsive voxels. Error bars indicate the 95% confidence intervals of the betweensubjects mean (n = 8).

classifier accuracy. The LGN is one example, in which the its BOLD response is not different between the stimuli, however, the stimuli can still be accurately classified (Fig. 3) showing that they have a differential representation. RSA (Fig. 5) shows that this is driven best by the response to the high temporal frequency stimulus. In hMT, the BOLD response is surprisingly strong to the RG stimulus, yet this is not differentially represented as shown in the RSA analyses (Figs. 5 and 9). Likewise, the strong BOLD response to the Ach P-type stimulus is maintained from the early to the ventral visual areas, yet its differential representation as shown in the RSAs becomes significantly weaker. Overall, variations in BOLD amplitude responses, be it across stimulus type, size or ROI do not readily predict classifier or RSA performance, demonstrating these as important approaches likely to yield new insights to stimulus coding.

4. Discussion

Using fMRI to understand how the human visual cortex encodes information typically involves identifying stimulus features that are differentially encoded across visual cortex. This approach has successfully identified areas with different specializations, but relies on univariate biases. Here we take a different approach by selecting very simple stimuli that are likely to elicit responses in all areas along the visual hierarchy. We used classification analyses and RSA to assess and compare the stimulus dimensions, alone or in combination, that are the most relevant for each visual area. We used four diagnostic stimulus types: the two chromatic stimuli represent the two cone-opponent dimensions of human colour vision whereas the two achromatic stimuli represent two extremes of spatio-temporal contrast: low-spatial, high-temporal and low-temporal, high-spatial. These four stimuli preferentially target the subcortical LGN pathways (P, M and K), allowing testable predictions for whether cortical responses are associated with specific subcortical inputs. The strength of the RSA approach is that it enables quantitative comparisons between different models of stimulus encoding, both within each area and between different areas, in order to characterize how response preferences evolve from LGN and across the cortex.

4.1. Feature biases do not reflect a segregation of magno-, parvo- and koniocellular signals

We found little evidence that cortical responses reflected the distinctive subcortical origins of the four stimulus types. Of the five models that could reflect a segregation of signals from a distinct subcortical pathway, only the Ach(M)/TF model performed well; it was one of the top-performing models in V3A/B and hMT, and also had significantly positive partial correlation (controlling for the Ach(P)/SF model) in areas V1, V2, and LO. This model may reflect differential responses to magnocellular and non-magnocellular signals or simply a neural code for temporal frequency. This is consistent with the known strength of the M-cell projection to the dorsal pathway (Maunsell et al., 1990; Nassi and Callaway, 2007, 2006). However, the other two models which performed well, Ach(P)/SF and Chr/Ach, cannot be accounted for by a segregation of inputs from different subcortical pathways. This lack of evidence for response segregation based on LGN origin is consistent with previous work showing that the segregated M and P channels from the LGN undergo significant combination in early visual cortex (Nassi and Callaway, 2007, 2006), and in the ventral cortex (Ferrera et al., 1994, 1992; Ninomiya et al., 2011), although some human imaging studies have supported a segregation of M and P functions in early cortex (Dumoulin et al., 2017; Tootell and Nasr, 2017). We note that across all areas, the poorest classifier performance was for discriminating between the two isoluminant chromatic stimuli (RG and BY), despite the fact that these stimuli were highly visible and evoked robust BOLD responses, consistent with evidence for combinations of these distinct (parvocellular and koniocellular) cone-opponent responses in V1 (Conway, 2001; de Valois et al., 2000; Goddard et al., 2010).

4.2. Gradients for chromaticity and spatial frequency along the ventral visual pathway

The best performing model in early visual cortex (V1–V3) was the Ach(P)/SF model, whereas this switched to the Chr/Ach model in the ventral areas hV4, VO1 and VO2. Across these areas there was also a trend for decreasing correlation with the Ach(M)/TF model, although correlation with this model was weaker overall in these areas. Notably, with the use of partial correlations to control for the interdependencies of the two achromatic models, correlations with these three models are independent of one another, hence an area's correlation with one model does not predict its correlation with the other. This means that the significant linear trends we found across the ventral visual pathway are independent: from V1 to higher ventral areas there is a decreasing correlation with the Ach(P)/SF and Ach(M)/TF models and an increasing correlation with the Chr/Ach model.

The relative importance of spatial form in driving responses in early visual cortex is consistent with area V1 having orderly maps for spatial frequency and orientation (Nauhaus et al., 2012). The very low correlation with the Chr/Ach model in V1 might reflect the fact that although a high proportion of V1 cells respond to colour, most of these also respond to achromatic form (Shapley and Hawken, 2011), mitigating against a strong segregation of chromatic and achromatic contrast responses. Any segregation of chromatic responses into cytochrome oxidase 'blobs' in primate V1, which remains controversial (Valverde Salzmann et al., 2012), is not reflected in the performance of our Chr/Ach model.

Although higher-level ventral areas still have a robust response to the high spatial frequency stimulus (Fig. 9), and the correlation with the Ach(P)/SF model remains positive, the magnitude of this correlation decreases along the ventral visual stream. This trend is consistent with the results of Salmela et al. (2016), who used a very different stimulus set but observed a similar trend for decreasing correlation with a spatial frequency model along the ventral visual stream. This might be related to areas such as hV4 and LO encoding higher-level form, including object shape (Cichy et al., 2011; DiCarlo and Cox, 2007; Oleskiw et al., 2018; Pasupathy and Connor, 2001), and becoming less sensitive to lower-level feature dimensions like spatial frequency. Across the ventral visual stream, receptive field sizes generally increase (Harvey and Dumoulin, 2011), and peak-SF tuning decreases (Henriksson et al., 2008) which could be related to the decreasing correlation with the Ach(P)/SF model. We think it unlikely that our result can be attributed to increasing receptive field size, not least because the higher-level ventral areas had robust responses to the high-SF stimulus (Fig. 9) as strong as to the other stimuli, and the high SF stimulus was above the peak-SF even for area V1 (Henriksson et al., 2008). Future work testing a larger range of spatial frequencies may reveal whether the higher ventral visual areas are less sensitive to spatial frequency overall and/or how this relates to shifts in their spatial frequency tuning.

The Chr/Ach model is defined by the prediction that responses to chromatic and achromatic stimuli are distinct from one another. It is not equivalent to a selective coding of stimulus colour, since it predicts that responses to achromatic and chromatic stimuli will be more discriminable than differences within these groups. The increasing correlation with the Chr/Ach model along the ventral visual stream suggests that achromatic and chromatic contrast information are carried by increasingly distinctive responses, perhaps by different subpopulations of neurons. Given the evidence in humans and non-human primates for a segregation of chromatic from achromatic contrast responses within the thin-stripes of V2 and V3, extending into V4, (Nasr et al., 2016; Tootell and Nasr, 2017; Tanigawa et al., 2010), it is interesting that the Chr/Ach model performs well in all these areas. Little is known about the organization of chromatic and achromatic responses in the VO areas, but these results suggest a differential segregation. In summary, our results show that, while the voxel patterns in the early visual areas have more differential information about achromatic form, voxel patterns in areas V4, VO1 and VO2 of the ventral pathway have relatively more information about chromatic/achromatic contrast. This is loosely supported by other RSA evidence demonstrating the importance of V4 for colour (Bannert and Bartels, 2018). In general, the heightened accuracy of chromatic vs achromatic classifier performance is compatible with specialization for colour processing developing in V4 and progressing through VO1 and VO2 of the human ventral visual cortex.

4.3. Implications for RSA for higher-level stimulus dimensions

Generating testable models of how complex responses emerge from the responses of earlier visual areas is a key goal of visual neuroscience

(DiCarlo et al., 2012). These gradients of response preferences along the ventral visual pathway are relevant to understanding the emergence of higher-level response properties (e.g. Salmela et al., 2016). The ventral visual stream responds to high-level form and is proposed to contain a 'object-form topography' (Haxby et al., 2001). An ongoing challenge is to characterize the extent to which there exist object-category-selective responses that cannot be accounted for by low-level visual similarity, since greater within-category than between-category visual similarity could yield an artefactual 'category' response (Andrews et al., 2015; Coggan et al., 2016; Rice et al., 2014; Watson et al., 2016). One approach (Connolly et al., 2012; Kriegeskorte et al., 2008), is to demonstrate categorical responses unique to higher areas. However, our results show that even for the simple stimuli used here there was substantial variation across visual cortex in which stimulus 'features' dominate the response. Simply demonstrating an area's response profile differs from that of early visual cortex is not sufficient to rule out a low-level explanation: gradients in responses to low-level stimulus properties mean that differences between V1 and later visual areas may reflect the emergence of a higher-order response or, instead, a reordered response to low-level features. This highlights the importance of generating targeted stimulus sets to disentangle low-level visual similarity driven responses from higher-level categorical responses (Bracci et al., 2017). A fuller characterization of response gradients for simple, easily-parameterized stimuli would also allow for better identification of stimulus dimensions that are relevant to higher-level areas (Goddard et al., 2018). In these ways, understanding how gradients in responses to low-level stimulus properties evolve across cortex is valuable for constructing models that bridge the gap between known tuning properties of early visual cortex and the emergence of more complex tuning.

5. Conclusions

We have used RSA and classification analyses to make novel quantitative comparisons across LGN and multiple visual areas of responses to low-level stimulus information. The simple stimuli we selected both preferentially target the subcortical LGN pathways that input visual cortex and are likely to elicit responses in all areas along the visual hierarchy. Our results show no evidence for response segregation based on LGN layer of origin. Instead, in the early visual areas (V1–V3), we find a processing bias with the strongest differential response for the achromatic, high spatial frequency stimuli, suitable for form vision, and some responsiveness to temporal frequency, but very little differential representation of chromatic vs achromatic contrast. In the ventral areas (V4, VO1 and VO2) we find the most differentiated representation is for chromatic vs achromatic contrast, implying a segregation of colour from achromatic contrast in these areas. In the dorsal areas, our results replicated the known response preferences for motion over form in achromatic contrast. Finally, our results also emphasize the importance of considering whether observed differences across visual areas could be accounted for by differences in responses to low-level stimulus properties, before inferring a more complex transformation of stimulus information.

CRediT authorship contribution statement

Erin Goddard: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft. **Kathy T. Mullen:** Conceptualization, Writing - original draft.

Acknowledgements

This work was funded by Canadian Institutes of Health Research (CIHR) grant 153277 and Natural Sciences and Engineering Research Council of Canada (grant RGPIN 183625-05) to KTM. We thank Michael Ferreira for his advice on fMRI protocols and the imaging technicians at the McConnell Brain Imaging Center (David Costa, Ron Lopez and Louise Marcotte) for their assistance with MRI data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2020.116780.

Data availability

Data from fMRI experiments are freely available online from the Open Science Framework (doi: https://doi.org/10.17605/OSF.IO/TF79N). This online repository includes deidentified raw data from the fMRI experiments, details of the stimulus timing for each participant, and the AFNI code used to perform the analyses reported here.

References

Allen, D., Hess, R.F., 1992. Is the visual field temporally homogeneous? Vis. Res. 32, 1075–1084. https://doi.org/10.1016/0042-6989(92)90008-7.
Anderson, S.J., Mullen, K.T., Hess, R.F., 1991. Human peripheral spatial resolution for achromatic and chromatic stimuli: limits imposed by optical and retinal factors. J. Physiol. 442, 47–64.
Andrews, T.J., Watson, D.M., Rice, G.E., Hartley, T., 2015. Low-level properties of natural images predict topographic patterns of neural response in the ventral visual pathway. J. Vis. 15, 3. https://doi.org/10.1167/15.7.3.
Bannert, M.M., Bartels, A., 2018. Human V4 activity patterns predict behavioral performance in imagery of object color. J. Neurosci. Off. J. Soc. Neurosci. 38, 3657–3668. https://doi.org/10.1523/JNEUROSCI.2307-17.2018.
Bannert, M.M., Bartels, A., 2017. Invariance of surface color representations across illuminant changes in the human cortex. Neuroimage 158, 356–370. https://doi.org/10.1016/j.neuroimage.2017.06.079.
Bannert, M.M., Bartels, A., 2013. Decoding the yellow of a gray banana. Curr. Biol. 23, 2268–2272. https://doi.org/10.1016/j.cub.2013.09.016.
Bird, C.M., Berens, S.C., Horner, A.J., Franklin, A., 2014. Categorical encoding of color in the brain. Proc. Natl. Acad. Sci. U.S.A. 111, 4590–4595. https://doi.org/10.1073/pnas.1315275111.
Borghesani, V., Pedregosa, F., Buiatti, M., Amadon, A., Eger, E., Piazza, M., 2016. Word meaning in the ventral visual path: a perceptual to conceptual gradient of semantic coding. Neuroimage 143, 128–140. https://doi.org/10.1016/j.neuroimage.2016.08.068.
Bracci, S., Ritchie, J.B., de Beeck, H.O., 2017. On the partnership between neural representations of object categories and visual features in the ventral visual pathway. Neuropsychologia 105, 153–164. https://doi.org/10.1016/j.neuropsychologia.2017.06.010.
Bradley, A., Zhang, X., Thibos, L., 1992. Failures of isoluminance caused by ocular chromatic aberrations. Appl. Optic. 31, 3657–3667.
Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vis. 10, 433–436.
Brewer, A.A., Liu, J., Wade, A.R., Wandell, B.A., 2005. Visual field maps and stimulus selectivity in human ventral occipital cortex. Nat. Neurosci. 8, 1102–1109.
Cichy, R.M., Chen, Y., Haynes, J.-D., 2011. Encoding the identity and location of objects in human LOC. Neuroimage 54, 2297–2307. https://doi.org/10.1016/j.neuroimage.2010.09.044.
Coggan, D.D., Liu, W., Baker, D.H., Andrews, T.J., 2016. Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. Neuroimage 135, 107–114. https://doi.org/10.1016/j.neuroimage.2016.04.060.
Connolly, A.C., Guntupalli, J.S., Gors, J., Hanke, M., Halchenko, Y.O., Wu, Y.-C., Abdi, H., Haxby, J.V., 2012. The representation of biological classes in the human brain. J. Neurosci. 32, 2608–2618. https://doi.org/10.1523/JNEUROSCI.5547-11.2012.
Conway, B.R., 2001. Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). J. Neurosci. 21, 2768–283.
Cottaris, N.P., 2003. Artifacts in spatiochromatic stimuli due to variations in preretinal absorption and axial chromatic aberration: implications for color physiology. J. Opt. Soc. Am. -Opt. Image Sci. Vis. 20, 1694–1713.
Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.
Dacey, D.M., 2000. Parallel pathways for spectral coding in primate retina. Annu. Rev. Neurosci. 23, 743–775.
Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I: segmentation and surface reconstruction. Neuroimage 9, 179–194.
de Valois, R.L., Cottaris, N.P., Elfar, S.D., Mahon, L.E., Wilson, J.A., 2000. Some transformations of color information from lateral geniculate nucleus to striate cortex. Proc. Natl. Acad. Sci. U.S.A. 97, 4997–5002.
DeSimone, K., Viviano, J.D., Schneider, K.A., 2015. Population receptive field estimation reveals new retinotopic maps in human subcortex. J. Neurosci. Off. J. Soc. Neurosci. 35, 9836–9847. https://doi.org/10.1523/JNEUROSCI.3840-14.2015.
DiCarlo, J.J., Cox, D.D., 2007. Untangling invariant object recognition. Trends Cognit. Sci. 11, 333–341. https://doi.org/10.1016/j.tics.2007.06.010.

DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object recognition? Neuron 73, 415–434. https://doi.org/10.1016/j.neuron.2012.01.010.
Dumoulin, S.O., Harvey, B.M., Fracasso, A., Zuiderbaan, W., Luijten, P.R., Wandell, B.A., Petridou, N., 2017. In vivo evidence of functional and anatomical stripe-based subdivisions in human V2 and V3. Sci. Rep. 7, 733. https://doi.org/10.1038/s41598-017-00634-6.
Dumoulin, S.O., Wandell, B.A., 2008. Population receptive field estimates in human visual cortex. Neuroimage 39, 647–660.
Engel, S.A., Rumelhart, D.E., Wandell, B.A., Lee, A.T., Glover, G.H., Chichilnisky, E.J., Shadlen, M.N., 1994. fMRI of human visual cortex. Nature 369, 525.
Farnsworth, D., 1957. The Farnsworth-Munsell 100-hue Test for the Examination of Color Discrimination.
Ferrera, V.P., Nealey, T.A., Maunsell, J.H., 1994. Responses in macaque visual area V4 following inactivation of the parvocellular and magnocellular LGN pathways. J. Neurosci. 14, 2080–2088.
Ferrera, V.P., Nealey, T.A., Maunsell, J.H., 1992. Mixed parvocellular and magnocellular geniculate signals in visual area V4. Nature 358, 756–761.
Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. Neuroimage 9, 195–207.
Goddard, E., Chang, D.H.F., Hess, R.F., Mullen, K.T., 2019. Color contrast adaptation: fMRI fails to predict behavioral adaptation. Neuroimage 201, 116032. https://doi.org/10.1016/j.neuroimage.2019.116032.
Goddard, E., Klein, C., Solomon, S.G., Hogendoorn, H., Carlson, T.A., 2018. Interpreting the dimensions of neural feature representations revealed by dimensionality reduction. Neuroimage 180, 41–67. https://doi.org/10.1016/j.neuroimage.2017.06.068.
Goddard, E., Mannion, D.J., McDonald, J.S., Solomon, S.G., Clifford, C.W.G., 2011. Color responsiveness argues against a dorsal component of human V4. J. Vis. 11 (3), 1–21. https://doi.org/10.1167/11.4.3.
Goddard, E., Mannion, D.M., McDonald, J.S., Solomon, S.G., Clifford, C.W.G., 2010. Combination of subcortical color channels in human visual cortex. J. Vis. 10 (25), 1–17. https://doi.org/10.1167/10.5.25.
Goddard, E., Mullen, K.T., 2019. fMRI responses to foveal versus peripheral chromatic and achromatic stimuli. J. Vis. 19 https://doi.org/10.1167/19.10.69, 69–69.
Goddard, E., Solomon, S.G., Carlson, T.A., 2017. Dynamic population codes of multiplexed stimulus features in primate area MT. J. Neurophysiol. 118, 203–218. https://doi.org/10.1152/jn.00954.2016.
Harvey, B.M., Dumoulin, S.O., 2011. The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture. J. Neurosci. Off. J. Soc. Neurosci. 31, 13604–13612. https://doi.org/10.1523/JNEUROSCI.2572-11.2011.
Haxby, J.V., Connolly, A.C., Guntupalli, J.S., 2014. Decoding neural representational spaces using multivariate pattern analysis. Annu. Rev. Neurosci. 37, 435–456. https://doi.org/10.1146/annurev-neuro-062012-170325.
Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430. https://doi.org/10.1126/science.1063736.
Henriksson, L., Nurminen, L., Hyvärinen, A., Vanni, S., 2008. Spatial frequency tuning in human retinotopic visual areas. J. Vis. 8 (5), 1–13.
Huk, A., Dougherty, R., Heeger, D., 2002. Retinotopy and functional subdivision of human areas MT and MST. J. Neurosci. 22, 7195–7205.
Ishihara, S., 1990. Ishihara's Tests for Color-Blindness, 38 Plate. Kanehara, Shuppan Co. Ltd., Tokyo/Kyoto.
Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8, 679–685.
Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.
Kleiner, M., Brainard, D., Pelli, D.G., 2007. What's New in Psychtoolbox-3? Perception, vol. 36. ECVP Abstract Supplement.
Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., Bandettini, P.A., 2008. Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60, 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043.
Lakens, D., 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front. Psychol. 4, 863. https://doi.org/10.3389/fpsyg.2013.00863.
Larsson, J., Heeger, D.J., 2006. Two retinotopic visual areas in human lateral occipital cortex. J. Neurosci. 26, 13128–13142.
Lee, B.B., Martin, P.R., Valberg, A., 1989. Nonlinear summation of M- and L-cone inputs to phasic retinal ganglion cells of the macaque. J. Neurosci. 9, 1433–1442.
Liu, J., Wandell, B.A., 2005. Specializations for chromatic and temporal signals in human visual cortex. J. Neurosci. 25, 3459–3468.
Martin, P.R., Lee, B.B., 2014. Distribution and specificity of S-cone ("blue cone") signals in subcortical visual pathways. Vis. Neurosci. 31, 177–187. https://doi.org/10.1017/S0952523813000631.
Maunsell, J.H., Nealey, T.A., DePriest, D.D., 1990. Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. J. Neurosci. Off. J. Soc. Neurosci. 10, 3323–3334.
Merigan, W.H., Katz, L.M., Maunsell, J.H., 1991. The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. J. Neurosci. 11, 994–1001.
Merigan, W.H., Maunsell, J.H., 1993. How parallel are the primate visual pathways? Annu. Rev. Neurosci. 16, 369–402. https://doi.org/10.1146/annurev.ne.16.030193.002101.
Merigan, W.H., Maunsell, J.H., 1990. Macaque vision after magnocellular lateral geniculate lesions. Vis. Neurosci. 5, 347–352.
Michna, M.L., Yoshizawa, T., Mullen, K.T., 2007. S-cone contributions to linear and non-linear motion processing. Vis. Res. 47, 1042–1054. https://doi.org/10.1016/j.visres.2007.01.014.
Mullen, K.T., 1985. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. J. Physiol. 359, 381–400.
Mullen, K.T., Chang, D.H.F., Hess, R.F., 2015. The selectivity of responses to red-green colour and achromatic contrast in the human visual cortex: an fMRI adaptation study. Eur. J. Neurosci. 42, 2923–2933. https://doi.org/10.1111/ejn.13090.
Mullen, K.T., Dumoulin, S.O., McMahon, K.L., de Zubicaray, G.I., Hess, R.F., 2007. Selectivity of human retinotopic visual cortex to S-cone-opponent, L/M-cone-opponent and achromatic stimulation. Eur. J. Neurosci. 25, 491–502.
Mullen, K.T., Kingdom, F.A.A., 2002. Differential distributions of red-green and blue-yellow cone opponency across the visual field. Vis. Neurosci. 19, 109–118.
Mullen, K.T., Thompson, B., Hess, R.F., 2010. Responses of the human visual cortex and LGN to achromatic and chromatic temporal modulations: an fMRI study. J. Vis. 10, 13. https://doi.org/10.1167/10.13.13.
Nassi, J.J., Callaway, E.M., 2007. Specialized circuits from primary visual cortex to V2 and area MT. Neuron 55, 799–808. https://doi.org/10.1016/j.neuron.2007.07.037.
Nasr, S., Polimeni, J.R., Tootell, Roger, 2016. Interdigitated color- and disparity-selective columns within human visual cortical areas V2 and V3. J. Neurosci. 36 (6), 1841–1857. https://doi.org/10.1523/JNEUROSCI.3518-15.2016.
Nassi, J.J., Callaway, E.M., 2006. Multiple circuits relaying primate parallel visual pathways to the middle temporal area. J. Neurosci. Off. J. Soc. Neurosci. 26, 12789–12798. https://doi.org/10.1523/JNEUROSCI.4044-06.2006.
Nauhaus, I., Nielsen, K.J., Disney, A.A., Callaway, E.M., 2012. Orthogonal micro-organization of orientation and spatial frequency in primate primary visual cortex. Nat. Neurosci. 15, 1683–1690. https://doi.org/10.1038/nn.3255.
Ninomiya, T., Sawamura, H., Inoue, K., Takada, M., 2011. Differential architecture of multisynaptic geniculo-cortical pathways to V4 and MT. Cerebr. Cortex 21, 2797–2808. https://doi.org/10.1093/cercor/bhr078.
Oleskiw, T.D., Nowack, A., Pasupathy, A., 2018. Joint coding of shape and blur in area V4. Nat. Commun. 9, 466. https://doi.org/10.1038/s41467-017-02438-8.
Pasupathy, A., Connor, C.E., 2001. Shape representation in area V4: position-specific tuning for boundary conformation. J. Neurophysiol. 86, 2505–2519.
Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vis. 10, 437–442.
Rice, G.E., Watson, D.M., Hartley, T., Andrews, T.J., 2014. Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. J. Neurosci. 34, 8837–8844. https://doi.org/10.1523/JNEUROSCI.5265-13.2014.
Saad, Z.S., Reynolds, R.C., Argall, B., Japee, S., Cox, R.W., 2004. SUMA: an interface for surface-based intra- and inter-subject analysis with AFNI. Proc IEEE Int. Symp. Biomed. Imaging Macro Nano 1510–1513.
Saad, Z.S., Ropella, K.M., Cox, R.W., DeYoe, E.A., 2001. Analysis and use of FMRI response delays. Hum. Brain Mapp. 13, 74–93.
Salmela, V.R., Henriksson, L., Vanni, S., 2016. Radial frequency analysis of contour shapes in the visual cortex. PLoS Comput. Biol. 12 https://doi.org/10.1371/journal.pcbi.1004719 e1004719.
Schiller, P.H., Logothetis, N.K., Charles, E.R., 1990. Functions of the colour-opponent and broad-band channels of the visual system. Nature 343, 68–70. https://doi.org/10.1038/343068a0.
Schira, M.M., Tyler, C.W., Breakspear, M., Spehar, B., 2009. The foveal confluence in human visual cortex. J. Neurosci. 29, 9050–9058.
Schneider, K.A., Richter, M.C., Kastner, S., 2004. Retinotopic organization and functional subdivisions of the human lateral geniculate nucleus: a high-resolution functional magnetic resonance imaging study. J. Neurosci. 24, 8975–8985.
Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893.
Shapley, R., Hawken, M.J., 2011. Color in the cortex: single- and double-opponent cells. Vis. Res. 51, 701–717. https://doi.org/10.1016/j.visres.2011.02.012.
Singh, K.D., Smith, A.T., Greenlee, M.W., 2000. Spatiotemporal frequency and direction sensitivities of human visual areas measured using fMRI. Neuroimage 12, 550–564.
Skerry, A.E., Saxe, R., 2015. Neural representations of emotion are organized around abstract event features. Curr. Biol. CB 25, 1945–1954. https://doi.org/10.1016/j.cub.2015.06.009.
Tanigawa, H., Lu, H.D., Roe, A.W., 2010. Functional organization for color and orientation in macaque V4. Nat. Neurosci. 13 (12), 1542–1548. https://doi.org/10.1038/nn.2676.
Tootell, R.B., Mendola, J.D., Hadjikhani, N.K., Ledden, P.J., Liu, A.K., Reppas, J.B., Sereno, M.I., Dale, A.M., 1997. Functional analysis of V3A and related areas in human visual cortex. J. Neurosci. 17, 7060–7078.
Tootell, R.B., Reppas, J.B., Kwong, K.K., Malach, R., Born, R.T., Brady, T.J., Rosen, B.R., Belliveau, J.W., 1995. Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. J. Neurosci. 15, 3215–3230.
Tootell, R.B.H., Nasr, S., 2017. Columnar segregation of magnocellular and parvocellular streams in human extrastriate cortex. J. Neurosci. Off. J. Soc. Neurosci. 37, 8014–8032. https://doi.org/10.1523/JNEUROSCI.0690-17.2017.
Valverde Salzmann, M.F., Bartels, A., Logothetis, N.K., Schüz, A., 2012. Color blobs in cortical areas V1 and V2 of the new world monkey Callithrix jacchus, revealed by non-differential optical imaging. J. Neurosci. Off. J. Soc. Neurosci. 32, 7881–7894. https://doi.org/10.1523/JNEUROSCI.4832-11.2012.

Vanni, S., Henriksson, L., Viikari, M., James, A.C., 2006. Retinotopic distribution of chromatic responses in human primary visual cortex. Eur. J. Neurosci. 24, 1821–1831.

Wardle, S.G., Kriegeskorte, N., Grootswagers, T., Khaligh-Razavi, S.-M., Carlson, T.A., 2016. Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. Neuroimage 132, 59–70. https://doi.org/10.1016/j.neuroimage.2016.02.019.

Watson, D.M., Young, A.W., Andrews, T.J., 2016. Spatial properties of objects predict patterns of neural response in the ventral visual pathway. Neuroimage 126, 173–183. https://doi.org/10.1016/j.neuroimage.2015.11.043.

Zhang, P., Zhou, H., Wen, W., He, S., 2015. Layer-specific response properties of the human lateral geniculate nucleus and superior colliculus. Neuroimage 111, 159–166. https://doi.org/10.1016/j.neuroimage.2015.02.025.