Gain control in the response of human visual cortex to plaids

J. Scott McDonald,1 Damien J. Mannion,2 and Colin W. G. Clifford3
1School of Psychology, The University of New South Wales, Sydney, New South Wales, Australia; 2Department of Psychology, University of Minnesota, Minneapolis, Minnesota; and 3School of Psychology and Australian Centre of Excellence in Vision Sciences, The University of Sydney, Sydney, New South Wales, Australia

Submitted 1 July 2011; accepted in final form 22 February 2012

McDonald JS, Mannion DJ, Clifford CWG. Gain control in the response of human visual cortex to plaids. J Neurophysiol 107: 2570–2580, 2012. First published February 29, 2012; doi:10.1152/jn.00616.2011.—A recent intrinsic signal optical imaging study in tree shrew showed, surprisingly, that the population response of V1 to plaid patterns comprising gratings components of equal contrast is predicted by the average of the responses to the individual components (MacEvoy SP, Tucker TR, Fitzpatrick D. Nat Neurosci 12: 637–645, 2009). This prompted us to compare responses to plaids and gratings in human visual cortex as a function of contrast and orientation. We found that the functional MRI (fMRI) blood oxygenation level-dependent (BOLD) responses of areas V1–V3 to a plaid comprising superposed gratings of equal contrast are significantly higher than the responses to a single component. Furthermore, the orientation response profile of a plaid is poorly predicted from a linear combination of the responses to its components. Together, these results indicate that the model of MacEvoy et al. (2009) cannot, without modification, account for the fMRI BOLD response to plaids in human visual cortex.

The dynamic response range of neurons is limited compared with the range of natural contrasts. Gain control allows neural sensitivity to be dynamically adjusted to suit the prevailing ambient contrasts (Crowder et al. 2008; Durant et al. 2007). This adjustment could be achieved through temporal sampling, i.e., adaptation (Gardner et al. 2005; Ohzawa et al. 1982), or spatial sampling in the area of the classic receptive field (Morrone et al. 1982) and extra classic receptive field (Levitt and Lund 1997) of a neuron. For two decades, divisive normalization has been the dominant model of contrast gain control (Brouwer and Heeger 2011; Heeger 1992). Principally intended to explain neuronal response saturation and cross-orientation suppression—suppression of response to optimally oriented stimuli by an overlaid stimulus of a different orientation—the original divisive normalization model proposed that neuronal response is inhibited by a large pool of neurons tuned to different orientations and spatial frequencies. Although subsequent research has challenged the notion that inhibitory neural circuitry is responsible for the divisive gain behavior (Carandini et al. 2002; Freeman et al. 2002; Li et al. 2006; Priebe and Ferster 2006), the model retains considerable descriptive power and attractive theoretical properties.

Two recent studies boast population, rather than isolated single-unit, measures of cross-orientation suppression (Busse et al. 2009; MacEvoy et al. 2009). Busse et al. (2009) found magnitudes of suppression in cat and human V1 consistent with traditional divisive normalization. However, MacEvoy et al. (2009) demonstrated that the response of tree shrew V1 to plaids is reliably predicted by the contrast-weighted average of the responses to the two components. In the case of two gratings of equal contrast, the cortical response is equal to the average of the responses to the individual components.

On the basis of their findings, MacEvoy et al. (2009) proposed that stimulus orientation is encoded through the relative activity of units tuned to different orientations, and the magnitude of the highest contrast component of the stimulus is encoded by the summed activity of the entire population. The underlying principle of this model is, hence, profoundly different from that of conventional divisive normalization models in which the summed activity of the entire population is used to normalize the response of individual neurons.

This leaves in question the nature of the normalization in humans, an important issue because humans are the species in which there is the greatest psychophysical characterization of contrast perception. Not only is the theoretical nature of cortical normalization in question, but also there is some 40% difference in the magnitude of response suppression observed in the studies of MacEvoy et al. (2009) and Busse et al. (2009). Therefore, we have investigated cross-orientation suppression in human visual cortex. We used the functional MRI (fMRI) blood oxygenation level-dependent (BOLD) signal to compare the response of early visual cortex to plaids with its response to the components of the plaids, first at a range of contrasts and then as a function of orientation.

METHODS

Experimental Procedures

The experimental protocols were approved by the Human Research Ethics Committee of the University of Sydney. Participants gave their written consent. Three women and five men took part in the first experiment. Four participants, one female, took part in the second experiment. All were experienced psychophysical observers and had normal or corrected-to-normal vision. Each participant completed ≥10 scans in which we measured cortical BOLD response while they viewed stimuli (henceforth functional runs): 4 or more runs to establish the retinotopic areas and 6 experimental runs.

Scanner and experimental setup. A Philips 3T scanner with a whole-head coil was used to conduct the MRI. Anatomical images (detailed recordings of the anatomy but with no functional information) were collected using a turbo field-echo protocol for enhanced gray-white contrast. They consisted of whole-head scans in the axial and sagittal planes (voxel size = 1.1 mm isotropic) and a high-resolution partial-head coronal scan (voxel size = 0.75 mm isotropic) to recover maximum detail in the occipital lobes, which contained the regions of specific interest to this study. Functional images (which are affected by the experimental manipulations) were collected using a SENSE (sensitivity encoding) sequence, a T2*-weighted, field-echo echo-planar imaging (boustrophedon) pulse sequence [repetition time (TR) = 3 s, echo time (TE) = 32 ms, flip angle = 90°, field of view...
where CMich is the Michelson contrast, Lmax is stimulus maximum (Yacoub et al. 2008). Each participant completed 4 runs for the grating luminance was equal to the mean luminance of the stimulus. The background was raised cosine window at its inner and outer edges. The background Eq. 1 presented at 40% contrast and the plaids at 80% contrast (as defined in sinusoidal gratings differing in orientation by 45°. The gratings were frequency of 3.35 cycles/° (Fig. 2) and plaids composed of two such marker presented at the center of the display. Stimuli were whenever there was a brief luminance decrement of the fixation blocks, participants were required to indicate with a button press to perform a demanding dimming task at fixation; throughout all 16th, and last blocks and 4 of the 8 experimental stimuli (each shown only) preceded all other stimulus conditions once. Individual runs consisted of 21 consecutive 15-s blocks of 5 different types (Fig. 1A). All runs contained fixation-only blocks, always the 1st, 6th, 11th, and last blocks and 4 of the 8 experimental stimuli (each shown 4 times within the run).

Each stimulus block contained 1 experimental stimulus presented at 20 different orientations over 15 s, each orientation shown only once, in pseudorandom order, for 0.75 s (see Fig. 1B). To equate apparent motion signals within the grating and plaid conditions, the order of the grating stimuli was constrained such that the change in orientation from 1 stimulus to the next was never >45°. Grating and plaid blocks were matched in the distribution of orientations to avoid potential biases introduced by anisotropies in response as a function of spatial orientation (Mannion et al. 2010; Sasaki et al. 2006).

Attention and eye movements were controlled by requiring observers to perform a demanding dimming task at fixation; throughout all blocks, participants were required to indicate with a button press whenever there was a brief luminance decrement of the fixation marker presented at the center of the display.

Second experiment. Stimuli were sinusoidal gratings with a spatial frequency of 3.35 cycles° (Fig. 2) and plaid composed of two such sinusoidal gratings differing in orientation by 45°. The gratings were presented at 40% contrast and the plaids at 80% contrast (as defined in Eq. 1). Both stimulus types were framed within an annulus of 0.7° inner radius and 7.2° outer radius. The annulus itself was enveloped by a 0.35° raised cosine window at its inner and outer edges. The background luminance was equal to the mean luminance of the stimulus.

We employed a continuous presentation paradigm to measure the BOLD response to stepwise rotation of the grating and plaid stimuli (Yacoub et al. 2008). Each participant completed 4 runs for the grating stimuli and 4 for the plaid stimuli, during which the stimulus orientation changed with each volume acquisition (3 s) in 11.25 shifts (see Fig. 2B). The direction of change alternated between clockwise and counterclockwise over runs. The stimuli were absent for the 1st and last 250 ms of each volume to prevent transients induced by abrupt changes in orientation. The spatial phase of the stimulus was randomly reassigned once every second. Stimulus onset, offset, and phase changes were presented in a square-wave temporal cycle. A full-orientation cycle was presented in 48 s (16 volumes). Each run consisted of 6 orientation cycles with a further 10 volumes (30 s) of blank fixation-only stimulation at both the beginning and end of the run.

Analysis Procedures

Preprocessing. A mean anatomic image was formed for each participant by combining the previously collected axial and sagittal whole-head scans and coronal partial-head scan. Before averaging, each anatomic image was inhomogeneity-corrected (Manjón et al. 2007), coregistered, and resampled to a voxel resolution of 0.75 mm (isotropic) where necessary. Each participant’s mean anatomic image was then segmented, voxels assigned as containing gray or white matter, using the automatic routines of mrGray (Tec et al. 1997) and ITKGray (Yushkevich et al. 2006; http://white.stanford.edu/software) followed by careful hand editing.

Between- and within-run participant movement was estimated and corrected by using the movement parameters estimated by SPM. The images were placed into register with the participant’s mean anatomic image by applying coregistration parameters to the affine transformation matrix of each image and reslicing using 4th degree B-spline interpolation; i.e., the scanned images were resampled so that they were congruent.

Regions of interest. For each participant, localizer scans using rotating wedge and expanding ring stimuli were employed to define regions of interest (ROIs) corresponding to the retinotopic visual areas V1, V2, V3, and hV4, following the definitions of Larsson and Heeger (2006). We defined hV4 as a hemifield representation of the contralateral visual field sharing a border with the ventral part of V3 and sharing the foveal representation of V1, V2, and V3 (Goddard et al. 2011).

Comparison of Conditions Using Percentage Signal Change

First experiment. For each participant, we extracted fMRI responses by averaging data from all the voxels within each ROI for each run and then averaged the signal across all blocks of the same type. The fMRI response in each condition was calculated as the percentage signal change (PSC) from fixation:

\[
PSC = 100 \times \frac{(t - b)}{b}
\]

where \( t \) is the mean signal value across the block (offset by 2 TRs to allow for hemodynamic delay) and \( b \) is the baseline response to the blank, fixation-only blocks. The PSC was then averaged across runs for each participant in each ROI.

Second experiment. We again extracted fMRI responses for each participant by averaging data from all the voxels within each ROI for each run. A correction of two volumes (6 s) was again applied to compensate for the lag in the hemodynamic response. Time courses from runs in which the stimulus advanced clockwise were temporally reversed and then combined with counterclockwise runs to produce mean time courses for stimulus orientation.

For each run, PSC was calculated for each orientation of the stimuli as defined in Eq. 2 for each cortical area. The baseline, \( b \), was taken to be the mean response to the first six volumes of the blank fixation-only stimulation (again offset by 6 s to compensate for hemodynamic lag) and the last six volumes. The mean signal value across all voxels was calculated for each orientation across the six complete orientation cycles in each run after discarding the response to the first quarter cycle of oriented stimulation. PSCs for each orientation were then averaged across runs.
RESULTS

First Experiment

We measured the BOLD response of retinotopically defined areas during the viewing of gratings and plaids of 10, 20, 40, and 80% Michelson contrast. The two components of each plaid pattern each had half of the Michelson contrast of the plaid pattern itself (i.e., component contrasts were 5, 10, 20, and 40%, respectively). We extracted the average time course across all voxels for each area and participant and calculated the PSC induced by the stimuli (see Eq. 2). As shown in Fig. 3, the measured PSC showed the expected increase with stimulus contrast for both gratings and plaids.

We compared the PSC induced by a grating of a given contrast with the PSC induced by a plaid with components of the same contrast as the grating (Fig. 4A). The findings of MacEvoy et al. (2009) suggest that the stimuli in this comparison would evoke equal magnitudes of cortical response. However, we find that the plaid PSC in visual areas V1, V2, and V3 is significantly greater than the PSC evoked by a grating with...
the same contrast as the plaid components (V1: $z = 3.88$, $P < 0.001$; V2: $z = 3.06$, $P = 0.002$; V3: $z = 3.06$, $P = 0.003$; 2-tailed Wilcoxon rank sign test). Only in hV4 is no significant difference observed ($z = 1.84$, $P = 0.110$).

We also find that the slope of the best-fitting line in the plaid and grating PSC comparison is steepest in V1 and declines in V2 and V3, reaching approximate unity in hV4 (Fig. 4B). However, comparing the relative PSC evoked by plaids and gratings in this way involves relating stimuli of differing Michelson contrast. When we investigate the relative PSC evoked by gratings and plaids of the same stimulus Michelson contrast, as shown in Fig. 4, C and D, such gratings and plaids produce magnitudes of PSC that are not significantly different in V1, V2, V3, and hV4 (all $Ps > 0.05$, 2-tailed Wilcoxon rank sign test). This suggests that Michelson contrast is the primary determinant of the magnitude of response to gratings and plaids.

Fig. 2. A: illustration of stimuli in experiment 2. A sinusoidal grating of 40% contrast (left) and 2 overlaid gratings (right) formed the plaid of 80% contrast. Both stimuli are shown at 0°. In the case of the plaid, this means that the 2 grating components were oriented at ±22.5° to the horizontal. Plaids contain regions where the bright and dark regions of the component gratings intersect and cancel out to midlevel gray. This modulation in contrast is a 2nd-order (non-Fourier) property of the stimulus as opposed to the 1st-order (Fourier) modulation of luminance in the component gratings. The orientation of these contrast modulations is perpendicular to that midway between the component gratings, and they are evident here as vertical gray stripes. B: a single cycle, 48-s long, of the stimulus presentation sequence. The stimulus orientation changes in 11.25° steps every 3 s. There is a 250-ms period at the onset and offset of an orientation when the stimulus is absent. On half of the runs, the grating orientation advanced counterclockwise (as shown here); on the other half, it advanced clockwise. C: schematics of the amplitude spectrum of the grating and plaid stimuli. $k$, Spatial frequency. The left represents a horizontal grating, and the middle represents a plaid with components at ±22.5° from horizontal. The right shows how the plaid is perceived; filled circles indicate a luminance modulation at the average of the 2 components, and the filled squares represent a contrast modulation orthogonal to the average orientation (Georgeson and Meese 1997).
Second Experiment

The results of the first experiment demonstrate that the average of the responses to the component gratings is a poor predictor of the response of human visual cortex to plaids. However, there are a number of plausible models described below that cannot be distinguished by these data. To address this issue, we conducted a second experiment comparing orientation response profiles for plaids with those for their component gratings.

We have recently demonstrated that several of the areas of human early visual cortex have anisotropic orientation tuning (Mannion et al. 2010; see also Saproo and Serences 2010; Swisher et al. 2010). The BOLD response in V1 is typically greatest at the obliques, weaker at vertical orientations, and weakest at horizontal orientations. This orientation response profile is a characteristic signature that allows us to compare the orientation response profile to a plaid pattern to that which would be predicted on the basis of the response to its component gratings in a way that would not be possible were the BOLD response isotropic as a function of stimulus orientation.

To exploit this feature of visual cortex, we measured the orientation tuning of V1, V2, V3, and hV4 with 40% contrast gratings and with 80% contrast plaids formed by linear superposition of 40% contrast gratings differing in orientation by 45°. We extracted the average time course for each area and participant and calculated the PSC induced by the stimuli across the series of orientations. The results, averaged across 4 participants, are shown in Fig. 5. The orientation of the plaid is defined as that midway between the two grating components. The cortical responses to gratings resemble the results of Mannion et al. (2010); in V1, the response to the gratings is greatest at obliques, weakest at horizontal, and intermediate for vertical orientations. Furthermore, in agreement with experiment 1, the plaid stimuli evoke a greater BOLD response than the gratings from all cortical areas.

We now compare these data to the predictions of four models (equations for each of the models are given in Table 1). The first “fixed-weight” model we test is that proposed by MacEvoy et al. (2009). The magnitude of the response to the plaid is the average of the responses to the two component gratings. The fixed-weight model’s prediction of the responses to the plaid stimuli (calculated for each participant and then averaged) and the actual responses (participants’ average) are compared in Fig. 6A. The model clearly underestimates the magnitude of the response to the plaids.

Comparison of the predictions of the fixed-weight model with the actual orientation response profile to plaids indicates that responses to plaid stimuli are not well-predicted by the contrast-weighted responses to the grating components. However, there might be other factors that affect cortical response and obscure the mechanism proposed by MacEvoy et al. (2009). One putative factor is attention. Despite the demanding attentional task performed at fixation by the participants during the scan, attentional enhancement could conceivably contribute more to the BOLD response for plaids than for gratings. Therefore, we test a second “fixed weights plus offset” model; BOLD responses to plaids are modeled as the average of the responses to the component gratings (as for the fixed-weight model) and an offset to represent any additional, attention-induced processing. The averages of the fits are shown in Fig. 6B. The magnitudes of the offsets, averaged across participants, for each area are plotted in Fig. 6C. We could reasonably expect the offset to be larger in hV4 than the earlier visual areas, as hV4 is subject to greater attentional modulation (Haenny and Schiller 1988; Hansen et al. 2007; Kastner et al. 1998; Maunsell 1995; Maunsell and Cook 2002; McAdams and Maunsell 1999; Moran and Desimone 1985; Schwartz et al. 2005; Tootell et al. 1998). In contrast, we find that the offset is smaller for hV4 than for the earlier visual areas. The fits of the fixed weights plus offset model are inevitably better than...
for the fixed weights model as the fixed weights model is essentially a reduced version of the fixed weights plus offset model with the offset set to 0. To compare such nested models, we follow Dobson (1990) in using an $F$ test to establish whether the fit of the more complex model is not only better, but also significantly better than that of simple one. If so, the introduction of the additional parameter is justified. If not, the simpler model is to be preferred. Comparison of residual errors show that the fit of the fixed weights plus offset model is significantly better than that of the fixed weights model in areas V1–V3 [V1: $F(1,14) = 13.5$, $P = 0.003$; V2: $F(1,14) = 12.2$, $P = 0.004$; V3: $F(1,14) = 7.2$, $P = 0.02$], justifying inclusion of the offset parameter for those areas, but misses significance in hV4 [$F(1,14) = 3.3$, $P = 0.09$].

In the third “free-weight” model, we allow a single free parameter to vary: the response to the plaid is the sum of the responses of the two gratings weighted by a parameter (which was fixed at 0.5 in the fixed-weight model). This can be regarded as a second test of the attention hypothesis; this time, however, the effect of attention on the BOLD response to plaids is multiplicative rather than additive. The data are plotted in Fig. 6D. The weights for each visual area, averaged across participants, are plotted in Fig. 6E. The quality of the fits is comparable with the offset model. Again, we find that the free parameter, which can be regarded as representing attentional amplification of the BOLD response, is actually lowest in hV4 and highest in V1. This is opposite to what we would expect of attentional modulations.

The results of the second experiment effectively rule out the application of a model of the form proposed by MacEvoy et al. (2009) to gain control in human visual cortex, even allowing for any potentially modulatory effects of attention on the response to plaid patterns. However, another difference between plaids and gratings is the existence of second-order, non-Fourier components in the plaid stimulus (Fig. 2). Wenderoth et al. (1999) found evidence suggesting that V3 is sensitive to the second-order components of plaids. More recently, we have demonstrated that V3, V2, and, to a lesser extent, V1 are all sensitive to the orientation of second-order contours defined by motion boundaries (Clifford et al. 2009), Larsson et al. (2006) have shown that V1, V2, V3, and hV4 are sensitive to second-order contrast and orientation modulation, and Hallum et al. (2011) have demonstrated that V1 is selective for second-order spatial frequency. We wish to investigate whether there is any way in which the original MacEvoy et al. (2009) model might be augmented to bring it into line with our data or whether they are fundamentally irreconcilable. Consequently, we formulate a simple elaboration of the free-weight model...
model that takes into account the existence of oriented second-order components in plaids. To the weighted sum of the gratings component responses, we add a third version of the same orientation response profile to represent the response to the second-order component of the plaid. This “non-Fourier” model has two free parameters: one weighting for the response to the two component gratings and the other weighting for the response to the second-order component.

There are two aspects of the non-Fourier model that require justification. First, we assume that the orientation response profile to second-order image structure is the same as that to first-order orientation. Little is known about the response of human visual cortex to second-order stimuli as a function of their orientation. Perhaps the most relevant finding is that the response of areas V1–V3 to motion-defined contours shows a similar pattern of bias for radial orientations (Clifford et al. 2009) to that reported earlier for luminance-defined stimuli (Sasaki et al. 2006). Here, we assume that the homology in the patterns of bias to the orientation of first- and second-order stimuli extends from meridian-relative to field-independent frames of reference. Second, we model the response to the salient second-order contrast modulation but do not try to model explicitly the response to the apparent modulation at the average of the two components. The justification for this is that it would be hard to separate with any confidence the response to the modulation from the response to the first-order components from which it differs in orientation by only $\pm 22.5^\circ$. Thus the proposed non-Fourier model should not be viewed as the only way in which the model of MacEvoy et al. (2009) might be augmented to incorporate sensitivity to second-order orientation.

The average of the participants’ fit of the non-Fourier model is plotted in Fig. 6F. The model captures the magnitude of the response at different orientations well and also accommodates responses at vertical and horizontal orientations better than the previous models. Critically, the weights for first-order components are not significantly different from the 0.5 predicted by MacEvoy et al. (2009) [V1: $t(3) = 1.11, P = 0.35$; V2: $t(3) = 0.21, P = 0.85$; V3: $t(3) = -0.57, P = 0.61$; hV4: $t(3) = -1.10, P = 0.35$]. To check that the superior fit is not simply because the model has more parameters, we used an $F$ test to test that the non-Fourier model performs significantly better than its counterparts. The non-Fourier model is significantly better than the free-weight model in all of the areas V1–hV4: V1 $F(1,14) = 19.8, P < 0.001$; V2 $F(1,14) = 42.6, P < 0.001$; V3 $F(1,14) = 55.9, P < 0.001$; hV4 $F(1,14) = 28.0, P < 0.001$. The non-Fourier model is also significantly better than the fixed weights plus offset model in areas V2–hV4, missing significance in V1 $F(1,14) = 3.2, P = 0.0940$; V2 $F(1,14) = 16.2, P < 0.001$; V3 $F(1,14) = 32.1, P < 0.001$; hV4 $F(1,14) = 20.1, P < 0.001$.

Interestingly, when we compare the weights for the gratings components and the second-order component in the non-Fourier model, a pattern emerges across the visual hierarchy (Fig. 6G). The weight for the gratings components is greatest in V1 and smallest in hV4. Conversely, the weight for the second-order component is greatest in hV4 and smallest in V1. Analysis of the weights across areas V1–hV4 using orthogonal polynomial contrasts (Winer et al. 1991) reveals significant linear trends [component gratings: $F(1,3) = 26.8, P = 0.02$; second-order component: $F(1,3) = 74.6, P = 0.004$] with no significant quadratic trend.

**DISCUSSION**

We measured the BOLD response to gratings and plaids in retinotopically defined areas of human visual cortex in two ex-
periments, first as a function of contrast and then as a function of orientation. We found that plaids evoked less response than the sum of their components but more than either individual component. We then exploited the presence of anisotropic BOLD responses as a function of grating orientation to compare a set of simple models and found that the plaid orientation response profile is best predicted by a weighted combination of responses to its first- and second-order (non-Fourier) components. The finding that the BOLD response to a plaid is consistently less than the sum of the responses to the components is
compatible with previous single-cell investigations into cross-orientation suppression in V1 (Bonds 1989; DeAngelis et al. 1992; Morrone et al. 1982) and the psychophysical finding that plaids have lower perceived contrast than gratings of the same Michelson contrast but greater perceived contrast than their constituent components (Georgeson and Shackleton 1994). In terms of magnitude of response, our data are also in good agreement with the human EEG results of Busse et al. (2009) and the fMRI data of Moradi and Heeger (2009). Both of these studies reported greater responses from a plaid than an individual component grating in human early visual cortex but less than linear summation. Our first experiment is also consistent with the PET data of Wenderoth et al. (1999), wherein plaids and gratings of similar physical contrast evoked similar response in human V1 and V2.

Although our results and those of Busse et al. (2009) and MacEvoy et al. (2009) all demonstrate that populations of neurons exhibit decreased responses to component gratings when they are part of a plaid, there is a large difference in the magnitude of suppression measured by MacEvoy et al. (2009) compared with our findings and those of Busse et al. (2009). The difference is exemplified in the results of our first experiment, shown in Fig. 3, wherein the ratios of plaid to grating responses in V1 tend to be ~30% greater than those found by MacEvoy et al. (2009). When we use the fixed-weight model, based on MacEvoy et al. (2009)’s proposed coding scheme, the predicted responses to the plaids fall well below the measured responses, as illustrated in Fig. 6A.

Although a comparison of our results with those of MacEvoy et al. (2009) is particularly pertinent because both involve measures based on cortical hemodynamic rather than direct neuronal activity, one significant methodological difference is that our participants were awake and thus subject to attention as a potential confounding factor. Despite participants performing a task designed to monopolize attentional processing, it remained plausible that the coding scheme in operation was essentially that proposed by MacEvoy et al. (2009) but obscured by enhanced attention-induced activity. However, when we fitted a model of fixed weights with additional offset to the results of the second experiment (illustrated in Fig. 6B), we found that V1 had the greatest offset and hV4 the smallest. Likewise, when we allowed the magnitude of the responses of the plaids to vary in the free-weight model (illustrated in Fig. 6D), we also found that in V1 the grating responses had to be modulated more than in hV4 to achieve the best fit. Given that hV4 is generally accepted to be more susceptible to attentional modulation than earlier visual areas (Haenny and Schiller 1988; Hansen et al. 2007; Kastner et al. 1998; Maunsell 1995; Maunsell and Cook 2002; McAdams and Maunsell 1999; Moran and Desimone 1985; Schwartz et al. 2005; Tootell et al. 1998), this is opposite to what would be expected if there were increased activity that corresponded to attentional processing.

The disparity between the studies could also be a result of the particular technique used: intrinsic signal optical imaging vs. fMRI of the T2*-weighted gradient-echo BOLD response. MacEvoy et al. (2009) describe their optical imaging data as predominantly reflecting blood-flow changes in cortical layers 2/3, and this is supported by their single-unit and intracellular recordings. The principal sources of the T2*-weighted gradient echo BOLD signal are the superficial layers of the cortex and the interface between gray-matter and cerebrospinal fluid (Goense and Logothetis 2006). Regardless of the exact weighting of the modulation of the BOLD signal across the layers of the cortex, we can assume that it reflects a wider breadth of cortical depths than optical imaging. The data reported here are then presumably more representative of the general cortical response.

Given that our data are not well-explained by the fixed-weight model of MacEvoy et al. (2009), we allowed the weighting of the component responses to vary for a new model of the second experiment. This approach is analogous to the “equal-weight” model used by Busse et al. (2009) to model their population responses, who found it worked well for gratings of equal contrast. Here, like MacEvoy et al. (2009), we use the raw grating responses and do not attempt to model the underlying orientation response profiles. In agreement with the first experiment, the fitted weights of the model suggest that responses of the components neither simply sum nor are averaged but instead lie between these two extremes. However, although the model captures the magnitude of the plaid response well and can be interpreted in the conventional contrast normalization framework, it poorly captures the variations in response magnitude across orientation.

Plaids contain not only the (1st-order) modulation of luminance considered thus far, but also a (2nd-order) modulation of local contrast. Regions of the plaids have low contrast, visible as oriented stripes across the plaid even though average luminance is identical to nearby high-contrast regions. As such second-order attributes have been demonstrated to drive cortical BOLD response in humans (Clifford et al. 2009; Hallum et al. 2011; Larsson et al. 2006), our final model (non-Fourier) introduced an additional weighted orientation response profile to capture the cortical activity evoked by this second-order component. This model gave significantly better fits to the data than the other models. In this non-Fourier model, the weighting of the responses to the component gratings, the first-order stimulus components, decreases through the visual hierarchy, whereas the weighting of the response to the second-order component increases. This pattern of increasing second-order contribution is consistent with the trend in adaptation to second-order stimuli measured by Larsson et al. (2006) and is suggestive of a transformation in the representation of the plaid as the cortical hierarchy is ascended, from primarily in terms of its grating components in V1 toward a representation closer to its perceptual interpretation in higher visual areas.

Importantly, the non-Fourier model also suggests that the magnitude of normalization for the first-order components of the stimulus may approximate that reported by MacEvoy et al. (2009). Despite the trend for the weighting of the grating components to decrease across the visual hierarchy, the weightings are generally close to 0.5. Such model weights suggest that, in the first experiment, the degree of normalization of the first-order components may have been obscured in the BOLD signal by the response to the second-order component. This observation could plausibly explain the apparent discrepancy between our data and those of MacEvoy et al. (2009). For example, in the species used by MacEvoy et al. (2009), the tree shrew, V1 may not contain neurons sensitive to second-order modulation, or it might be that the response layers 2/3 of tree shrew, the primary contributors to the optical imaging signal, only reflect first-order processing even if it is present in deeper cortical layers. If this were true, it could also
explain the discrepancy between the results of MacEvoy et al. (2009) and the single-unit data recorded from cat V1 by Busse et al. (2009) as there is evidence of neurons in cat V1 that are responsive to second-order stimuli (Baker and Mareschal 2001). We have stressed the importance of second-order neural assemblies contributing to the BOLD response because of the ready evidence from previous fMRI studies (Clifford et al. 2009; Hallum et al. 2011; Larsson et al. 2006) and because of the compelling perceptual appearance of low-contrast contours in the stimuli. However, we note that BOLD responses may also be driven by neural assemblies tuned to other spatial properties. One possibility is the existence of explicit “plaid detectors” as suggested by recent psychophysical studies (McGovern and Peirce 2010; Peirce and Taylor 2006; Robinson and MacLeod 2011). If there are such mechanisms, supported perhaps by neurons like those described by Anzai et al. (2007) tuned to combinations of orientations, then these would also be expected to contribute to the BOLD response to plaids. Furthermore, whereas gratings have only one perceptual interpretation, plaids have two: they can be perceived as either two overlaid transparent gratings or as a checkerboard (Georgeson and Meese 1997, 1999, Meese and Georgeson 1996, 2005). Whatever mechanisms mediate the transparency/ checkerboard dominance, it is clear that the percept of plaids is determined by computations that do not accompany grating processing in the visual system. These mechanisms presumably also contribute to greater BOLD response to plaids compared with gratings.

In summary, the empirical findings of this study indicate that the model of MacEvoy et al. (2009) cannot, without modification, account for the fMRI BOLD response to plaids in human visual cortex. Modeling in terms of first- and second-order components suggests that our results may be broadly reconcilable with the model of MacEvoy et al. (2009) in terms of the magnitude of normalization of first-order stimulus components. However, our results show that their proposed contrast coding scheme cannot be applied as a general mechanism across all cortical areas. MacEvoy et al. (2009)’s model of contrast coding in V1 critically depends on contrast-weighted averaging of component responses in response to a plaid. Our best-fitting model involves significant change in this weighting across the cortical hierarchy, suggesting that not all visual areas can be following such a coding scheme. Future work will be required to test the validity of our best-fitting model and to provide a definitive model of contrast gain control in the response of human visual cortex to plaids and other more complex stimuli.

GRANTS

This work was supported by National Health and Medical Research Council of Australia Project Grants 570855 and 1027258 and an Australian Research Council Future Fellowship to C. W. G. Clifford.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


REFERENCES


