Orientation bandwidths are invariant across spatiotemporal frequency after isotropic components are removed

John Cass  
School of Psychology, University of Sydney, Australia

Sjoerd Stuit  
Helmholtz Institute, Department of Psychology, Universiteit Utrecht, Netherlands

Peter Bex  
Schepens Eye Research Institute, Harvard Medical School, USA

David Alais  
School of Psychology, University of Sydney, Australia

It is well established that mammalian visual cortex possesses a large proportion of orientation-selective neurons. Attempts to measure the bandwidth of these mechanisms psychophysically have yielded highly variable results (~6°–180°). Two stimulus factors have been proposed to account for this variability: spatial and temporal frequency; with several studies indicating broader bandwidths at low spatial and high temporal frequencies. We estimated orientation bandwidths using a classic overlay masking paradigm across a range of spatiotemporal frequencies (0.5, 2, and 8 c.p.d.; 1.6 and 12.5 Hz) with target and mask presented either monoptically or dichoptically. A standard three-parameter Gaussian model (amplitude and width, mean fixed at 0°) confirms that bandwidths generally increase at low spatial and high temporal frequencies. When incorporating an additional orientation-untilted (isotropic) amplitude component, however, we find that not only are the amplitudes of isotropic and orientation-tuned components highly dependent upon stimulus spatiotemporal frequency, but orientation bandwidths are highly invariant (~30° half width half amplitude). These results suggest that previously reported spatiotemporally contingent bandwidth effects may have confounded bandwidth with isotropic (so-called cross-orientation) masking. Interestingly, the magnitudes of all monoptically derived parameter estimates were found to transfer dichoptically suggesting a cortical locus for both isotropic and orientation-tuned masking.

Keywords: orientation, bandwidth, cross-orientation masking, spatial frequency, temporal frequency


Introduction

Orientation-selective neurons are common in primary visual cortex (V1) and are considered critical for edge extraction and image segregation (Hogan, Garraghty, & Williams, 1999; Hubel & Wiesel, 1962, 1968; Niell & Stryker, 2008; Ringach, Shapley, & Hawken, 2002; Usrey, Sceniak, & Chapman, 2003). However, orientation bandwidth estimates are surprisingly variable in their neural (2°–360°) and psychophysical (7°–360°) expression (Anderson & Burr, 1985; Anderson, Burr, & Morrone, 1991; Blake & Hologpian, 1985; Blakemore & Campbell, 1969; Campbell & Kulikowski, 1966; Movshon & Blakemore, 1973).

Variation in bandwidth over spatiotemporal frequency may account for this. Several psychophysical masking and adaptation studies indicate that bandwidths decrease with spatial frequency (Anderson & Burr, 1985; Blake & Hologpian, 1985; Kelly & Burbeck, 1987; Phillips & Wilson, 1984; Snowden, 1992) but increase with the temporal frequency of the stimulus (Anderson et al., 1991; Kelly & Burbeck, 1987; Snowden, 1992). Not all studies find this dependency on either spatial (Movshon & Blakemore, 1973) or temporal (Phillips & Wilson, 1984) frequency. Nonetheless a meta-analysis of earlier psychophysical studies reveals a significant negative correlation between orientation bandwidth and spatial frequency. The much weaker negative correlation between orientation bandwidth and temporal frequency is insignificant (see Figure 1). In contrast, neurophysiological evidence suggests little or no frequency dependency. A small fraction of V1 units show bandwidth broadening at low spatial frequencies (Mazer, Vinje, McDermott, Schiller, & Gallant, 2002). No temporal frequency dependency has been reported at population or single-cell levels (Moore, Alitto, & Usrey, 2005).

Stimulus and procedural differences between these studies may explain many of these apparently conflicting estimates of bandwidth. Electrophysiological studies derive estimates...
Figure 1. Meta-analysis of the relationship between: (a) spatial frequency and (b) temporal frequency; and orientation bandwidth (HWHA) as estimated by previous psychophysical studies (see inset for references). Linear pairwise correlation between orientation bandwidth and: spatial frequency (controlling for temporal frequency) ($r^2 = .490$, $p < .001$); and temporal frequency (controlling for spatial frequency) ($r^2 = .256$, $p = .07$). Open symbols signify the results of masking experiments; closed symbols, adaptation.
using single stimuli presented at various orientations, while masking/adaptation studies combine a fixed and a variable orientation. Psychophysically, it is known that stimuli beyond the target channel’s orientation bandwidth can alter threshold measurements (Boynton & Foley, 1999; Cass & Alais, 2006; Meese & Holmes, 2007; Meier & Carandini, 2002; Morrone, Burr, & Maffei, 1982). Similarly, in cross-orientation masking (XOM) (also known as cross-orientation suppression (XOS)), an optimally stimulated V1 neuron is suppressed by superimposed orientations beyond that neuron’s bandwidth (Allison, Smith, & Bonds, 2001; Bauman & Bonds, 1991; Carandini, Heeger, & Movshon, 1997; DeAngelis, Robson, Ohzawa, & Freeman, 1992; Freeman, Durand, Kiper, & Carandini, 2002; Morrone et al., 1982; Morrone, Burr, & Speed, 1987; Priebe & Ferster, 2006). It is unclear if psychophysical and physiological XOM are related, but both are biased towards low spatial and high temporal frequencies (Alison et al., 2001; Bauman & Bonds, 1991; Boynton & Foley, 1999; Cass & Alais, 2006; Meese & Hess, 2004; Meese & Holmes, 2007; Meier & Carandini, 2002), where psychophysical orientation bandwidth estimates are broadest (Anderson & Burr, 1985; Anderson et al., 1991; Phillips & Wilson, 1984; Sharpe & Tolhurst, 1973; Snowden, 1992).

In these experiments, we differentiate orientation-tuned from orientation-untuned (isotropic) masking. Two recent psychophysical studies, one employing overlay masking (Baker & Meese, 2007) and the other, temporal reverse psychophysical studies, one employing overlay masking from orientation-untuned (isotropic) masking. Two recent correlation (Roeber, Wong, & Freeman, 2008), found bandwidth estimates of approximately 20

\[
\text{sd} = 20 \text{ pixels}
\]

able from orientation-tuned masking, yielding orientation evidence for strong isotropic suppression that was dissociated in that XOM may be intra-cortical process.

**Stimuli**

The stimuli were spatiotemporally narrowband achromatic noise sequences presented within a circular aperture (diameter = 5 degrees of visual angle) with a cosine-ramped outer edge (ramp SD = 20 pixels) against a gray background held at mean luminance. Target and masking stimuli were each created from independent noise sources on each trial. Noise was generated by assigning each pixel within a 128 × 128 × 64 matrix, a luminance value derived from a uniformly random distribution of values between −1 and 1. The dc component of this distribution was set to zero and later rescaled to mean luminance (52 cd/m²) to ensure that the mean luminance of each image sequence was identical. The fast Fourier transform (FFT) was calculated for each image within each movie sequence image (i.e., spatially) and between images within each sequence (i.e., temporally). The spatiotemporal amplitude spectrum of each image was band-pass (SD = 0.1 octaves) centered at one of three spatial frequencies (0.5, 2 and 8 c.p.d.) and one of two temporal frequencies (1.6 and 12.5 Hz). For any given block of trials, the spatial and temporal frequency of target and masking stimuli were identical. The spatial and temporal phases of target and masking stimuli varied randomly both with respect to each other and across trials. Root mean squared (RMS) calculation of spatiotemporal image contrast was to constrain the luminance distribution of target and masking stimuli.

All target and masking stimuli were orientation filtered (SD = 1°). Target orientation was centered at 135° and masking orientations varied between 135° and 225°, producing a range of orientation difference between target and masking stimuli between 0° and 90° (orientation differences used: 0°, 2.5°, 5°, 6.125°, 10°, 12.5°, 20°, 30°, 45°, 67.5°, and 90°). The total duration of each stimulus pattern was 640 ms and was ramped on and off using a raised cosine (SD of ramp = 20 ms).

On any given trial, a single fusion-locking square (5 × 5 degrees of visual angle) defined by a 2 pixel wide black line appeared in each quadrant of the screen (total of four fusion locks). The center of each fusion lock was located 2.75 degrees above or below the horizontal midline and 6.25 degrees left or right of the vertical midline. Fixation squares (0.5 × 0.5 degrees) of uniform color (black, white, green, or red) were located on the horizontal midline and centered beneath each fusion lock. Target and masking stimuli were centered within fusion lock boundaries.

In monoptic viewing conditions, the target stimulus was spatially superimposed (added) onto one of the two

**Methods**

**Materials**

Experiments were programmed using the MATLAB (Mathworks Ltd) programming environment using programmes from the Psychophysics Toolbox version 3 (Brainard, 1997). Stimulus presentation was driven by an ATI Radeon X1600 graphics card installed in a Mac Pro quad core computer. Stimuli were displayed on a Mitsubishi Diamond Pro CRT Monitor (1024 × 768 pixel resolution in a display area of 24 × 24 cm, 100 Hz vertical refresh, mean luminance = 52 cd/m²) with a linearized gamma. The 10.8-bit luminance resolution was achieved using bit-stealing. Stimuli were viewed through a bench-mounted mirror stereoscope to bring the eyes into alignment. Total viewing distance (including reflective path) was 57 cm.
masking stimuli above or below fixation (see Figure 2a). In dichoptic viewing conditions, again target stimuli were presented either above or below fixation, but in this case were presented to the eye contralateral to that projected by the mask (see Figure 2b). The eye in which target and masking stimuli were presented was randomized across trials.

**Procedure**

The first experimental phase involved measuring target detection thresholds for each spatial and temporal frequency in the absence of masking stimuli. In the second stage, target detection thresholds were measured in the context of perceptually superimposed (physically superimposed in the case of monoptic presentation) with one of two identical masking stimuli whose contrast was fixed at 16 × RMS contrast threshold. The orientation of the target stimulus was fixed at 135°. The orientation of the masking stimulus and ocular viewing condition (monoptic vs. dichoptic) varied randomly across trials. The spatial and temporal frequencies of target and masking stimuli were blocked across trials and were chosen pseudorandomly.

A spatial two alternative forced choice procedure was used to estimate threshold. On any given trial the target stimulus appeared above or below fixation to either the left or right eye. Subjects were required to report via a keypress whether the target had been presented above or below fixation. Correct

![Figure 2. Schematic depiction of (a) monoptic and (b) dichoptic stimulus viewing conditions. Under monoptic conditions, the target and a single mask are physically superimposed. In dichoptic conditions, the target and mask are perceptually co-localized but presented to different eyes. Targets and masks presented above and below fixation could appear in either the same or different eyes. In experimental contexts, masking stimuli were presented at 16× detection threshold. Target contrasts appear in this figure at suprathreshold levels for illustrative purposes.](http://jov.arvojournals.org/)
and incorrect responses elicited subsequent decreases and increases in target contrast respectively. An adaptive staircase procedure was used to converge upon target detection threshold (40 trials/staircase). At least four separate runs were used to calculate thresholds. Psychometric functions were derived for each subject’s data pooled across separate runs using a bootstrapping algorithm with a lower limit set at 50% correct performance and target detection threshold estimates corresponding to 75% correct performance.

Corrective feedback was provided following each trial by briefly changing the color of the fixation point (red = incorrect; green = correct) immediately following the response. Color-cued performance feedback lasted for 500 ms. The luminance of the fixation point provided temporal information about the trial presentation. Prior to the onset of each trial, the fixation point was colored white (500 ms), then black throughout the duration of the trial (640 ms) followed by white prior to the subject’s response.

Figure 3. Effects of relative target-mask orientation on target detection threshold elevation at different spatial and temporal frequencies measured for three observers under monoptic and dichoptic masking conditions. The ordinate in each graph represents elevations in target detection threshold relative to baseline (no masking) conditions (dB = 20 × log 10 [masked threshold / baseline threshold]) and the abscissa, the difference in orientation between target and masking stimuli. Left-hand columns (red circles) represent monoptic masking conditions and right-hand columns (blue diamonds), dichoptic masking conditions. Each set of six graphs (defined by the borders of the gray background) represents three different subject’s data (rows) with each set corresponding to a different stimulus spatiotemporal frequency: sets from left to right: 0.5, 2, and 8 c.p.d.; top set: 1.6 Hz; bottom set: 12.5 Hz. Fits in each graph correspond to a standard Gaussian (dashed curve; see equation in Figure 4a) and a modified Gaussian with additional vertical offset (isotropic amplitude) parameter (solid curve; see equation in Figure 4b). Shaded regions in each graph correspond to the isotropic amplitude estimate (σ).
Observers

Three psychophysical observers with normal or corrected to normal vision participated in the experiment. All were right eye dominant as measured by a hole-in-the-card test. All were experienced psychophysical observers. One was naïve to the experiment’s purposes. The others were authors and were aware of the purposes of the experiment.

Results

Changes in target detection thresholds were measured as a function of the relative orientation difference between spatiotemporally overlayed target and masking stimuli, both modulated at a spatial frequency of 0.5, 2, or 8 c.p.d. and temporal frequency of 1.6 or 12.5 Hz. As can be seen in Figure 3, robust threshold elevations are observed at all spatiotemporal frequencies using both monoptic and dichoptic presentation, with little evidence of facilitation (threshold reduction). Figure 3 also demonstrates that in all spatiotemporal conditions the magnitude of threshold elevation is highly dependent upon the relative orientation of target and mask, with greater difference in orientation associated with a generally monotonic reduction in threshold.

Bandwidth ($\sigma$)

In order to infer the bandwidth of this relationship between threshold and target-mask orientation difference, we fit two separate functions to each subject’s data derived at each stimulus spatial and temporal frequency. The first fit, conforming to a standard Gaussian function (see Equation in Figure 4a), is composed of two free parameters (amplitude (A) and bandwidth ($\sigma$)) and a peak target-mask orientation difference parameter fixed at 0° (see dashed red and blue curves in Figure 3). The bandwidth estimates derived from this standard Gaussian fit are shown for each spatiotemporal frequency in Figure 5a. A three-way within-subjects ANOVA provides no evidence for any interaction between the effects of stimulus temporal frequency, spatial frequency, and ocularity (monoptic vs. dichoptic) on orientation bandwidths ($p > .05$). No significant two-way interactions were

$$G_1 = A \cdot \exp \left[ -\frac{(x-0)^2}{2 \cdot \sigma^2} \right]$$

$$G_2 = A \cdot \exp \left[ -\frac{(x-0)^2}{2 \cdot \sigma^2} \right] + \alpha$$

$$G_2 = G_1 + \alpha \left[ -\frac{(x-0)^3}{2 \cdot \sigma^2} \right] + \alpha$$

Figure 4. Visual and symbolic depiction of Gaussian models used to fit masking data in Figure 3. (a) Standard Gaussian model consisting of two free parameters: orientation-tuned amplitude (A) and orientation bandwidth ($\sigma$). (b) Modified Gaussian containing an additional free parameter: isotropic amplitude ($\alpha$). In both models, peak amplitude is fixed at 0° orientation difference.
observed between ocular presentation mode and spatial frequency or temporal frequency (both p-values > .05). Significant two-way interactions were observed, however, between spatial and temporal frequency when collapsing across monoptic and dichoptic conditions (F(2, 2) = 3.276, p < .05). As can be seen in Figure 5, this interaction is due to broader bandwidths in response to 12.5 Hz compared with 1.6 Hz modulation at the lower spatial frequencies tested (117.4° vs. 58.3° (0.5 c.p.d.; 96.4° vs. 36.7° (2 c.p.d.)). Bandwidths did not vary as a function of temporal frequency at the highest spatial frequency tested. In the case of 12.5 Hz modulation, 8 c.p.d. stimuli generated significantly narrower bandwidths than lower spatial frequencies. For 1.6 Hz stimuli, both higher spatial frequencies (2 and 8 c.p.d.) elicited narrower bandwidths than the lowest spatial frequency tested (0.5 c.p.d.).

To determine whether these spatiotemporal frequency-contingent bandwidth estimates are confounded by isotropic (cross-orientation) masking, we fitted these data with a modified Gaussian (see solid red and blue curves in Figure 3), which included an additional additive isotropic component (α) as a free parameter (see Equation 2, Figure 4b). Unsurprisingly, the addition of this isotropic free parameter improved the fits (average chi-square = 33.34 vs. 17.38). It is worth noting, however, that a within-subjects t-test comparing these chi-squared estimates indicates that the benefit imparted by this isotropic component is highly significant (t(22) = 3.56; p < .01).

As can be clearly observed in Figure 5b, incorporating this isotropic component has a striking effect on bandwidth estimates (compare dashed and solid curves). Not only are average bandwidth estimates far narrower on average (range = 22.6°–30.9°) than those derived using the standard Gaussian fit, but a three-way within-subjects ANOVA finds no significant interactions or main effects in bandwidth estimates as they occur across different spatial frequencies, temporal frequencies, or ocular modes of presentation (all p-values > .05).

Figure 6 shows the magnitude of the two free amplitude parameters derived from the modified Gaussian curve fits depicted in Figure 4b (orientation-tuned amplitude (A) and isotropic amplitude (α)) expressed as a function of the spatial and temporal frequency of target and masking stimuli. The left-hand column in Figure 6 represents parameter estimates as they occur monoptically (red) and dichoptically (blue), grouped by stimulus spatial and temporal frequency. The parameter estimates in the right column in Figure 6 show the same data collapsed across monoptic and dichoptic conditions, grouped by temporal frequency (x-axis) with gray levels representing different spatial frequency conditions.

Isotropic amplitude (α)

A three-way within-subjects ANOVA shows no significant interactions between ocular mode of presentation, spatial frequency, or temporal frequency on isotropic amplitude (α) (all p-values > .05). Significant differences in the amplitudes of isotropic masking components were observed for different spatial frequencies (F(2, 2) = 25.848, p < .05). Figure 6 reveals that this main effect is driven by higher isotropic masking amplitudes at the lowest spatial frequency tested, which is evident in both 1.6 and 12.5 Hz conditions. No significant differences in isotropic amplitude were observed as a function of spatial frequency or ocular mode of presentation (both p-values > .05).

Orientation-tuned Gaussian amplitude (A)

Separate within-subjects ANOVAs show that the amplitude of the orientation-tuned masking component depends neither upon both the spatial frequency nor the temporal frequency of the stimulus (p-values < .05).

Integral

The results above demonstrate that the magnitude of overlay masking is composed of both isotropic (α) and orientation-tuned (A) components, each of which exhibit
distinctive systematic contingencies across different regions of spatiotemporal frequency space. Each of these components ($\alpha$ and $A$) may be independently integrated as a function of target-mask orientation difference ($\int_0^{90} \alpha$ and $\int_0^{90} A$) to reveal its relative contribution the total masking effect ($\int_0^{90} \alpha + A$) as function of spatiotemporal frequency (collapsing across monoptic and dichoptic viewing conditions). As can be seen in Figure 7, the combined integral ($\int_0^{90} [\alpha + A]$) is maximal at 0.5 c.p.d. and decreases (approximately log-linearly) with spatial frequency. No significant differences in the combined integral are evident across temporal frequencies.
Separate within-subjects ANOVAs indicate no interaction at either temporal frequency between combined and isotropic integrals as a function of spatial frequency ($p > .05$). However, significant interactions are observed (for each temporal frequency) between combined and orientation-tuned integrals as a function of spatial frequency. These results imply that the spatial frequency dependencies of isotropic masking resemble more closely the combined isotropic and orientation-tuned masking effects than do the orientation-tuned masking effects. Figure 7 shows that this difference in the spatial frequency dependencies of isotropic and orientation-tuned integrals are due to significantly higher isotropic masking integrals at 0.5 c.p.d., with no differences observed at either 2 or 8 c.p.d..

**Discussion**

We employed a classic overlay masking paradigm in which we varied the relative orientation of target and masking stimuli to estimate the bandwidths of psychophysical orientation channels across various regions of spatiotemporal frequency space. When fitted with a standard Gaussian function (Equation 1, Figure 1a) (without an additive isotropic masking component), we found bandwidths to be highly contingent upon the spatiotemporal frequency of the stimulus. These bandwidth effects may be summarized as being broader at low spatial (0.5 c.p.d.) (using 1.6 Hz modulation) and high temporal frequencies (12.5 Hz). This dependency between spatiotemporal frequency and orientation bandwidth resembles several previous findings. For example, Anderson et al. (1991) and Snowden (1991) found that bandwidths are broader at low spatial and high temporal frequencies. Similarly, Phillips and Wilson (1984) reported that bandwidths decrease with increasing spatial frequency, although temporal frequency was found to have no effect.

We find that incorporating an isotropic amplitude component into our Gaussian fits abolishes all spatiotemporal frequency-dependent bandwidth effects. A similar abolition of spatial frequency-contingent bandwidth effects was reported by Movshon and Blakemore (1973) after correcting for adaptation-induced elevations in threshold. It is worth noting, however, that Movshon and Blakemore’s bandwidth estimates were somewhat narrower than ours ($\sim 6^\circ$–$11^\circ$ vs. $\sim 22^\circ$–$35^\circ$), with our estimates corresponding more closely to those derived from the masking study of Phillips and Wilson (1984) ($\sim 18^\circ$–$28^\circ$) (although we show no systematic variation with spatial frequency). We suggest that the differences in the magnitude of these bandwidth estimates point to a potential difference between the mechanisms of adaptation and masking.

That we observe no systematic co-variation in orientation bandwidth with spatial or temporal frequency suggests...
that earlier studies may have conflated within-orientation channel masking with isotropic (cross-orientation) masking (or adaptation). This conclusion is strengthened by the strong dependency of isotropic amplitude upon stimulus spatiotemporal frequency. Specifically, the lowest spatial frequency used in our study (0.5 c.p.d.) is associated with significantly higher isotropic amplitudes than the higher spatial frequencies tested (2 and 8 c.p.d.). That low spatial frequencies are capable of eliciting such narrow bandwidths in combination with strong isotropic masking has been reported previously in abstract form (Meese & Holmes, 2003, in press).

This low spatial frequency bias in isotropic masking squares with other psychophysical evidence indicating that XOM is strongly dependent on spatial frequencies ≤2 c.p.d. (Meese & Holmes, 2007, in press). In addition to the strong low spatial frequency dependence of XOM, these earlier studies found XOM to be contingent upon temporal frequencies ≥4 Hz. Our isotropic masking component does not share this spatiotemporal dependency. Rather, we find this isotropic component to be very strong at both 1.6 and 12.5 Hz at 0.5 c.p.d. Why this difference in the temporal tuning characteristics of isotropic masking should be evident between studies is uncertain. One factor may be that the low temporal frequency used in our study (1.6 Hz) modulated at approximately three times the low frequency modulation rate used by Meese and Holmes (2007) (nominally 0.5 Hz). It is possible, therefore, that our “low” temporal frequency (1.6 Hz) exceeded the lower temporal limits of XOM, below which the effect may cease to occur. We tested this in a single subject (author JC) using 0.5 c.p.d. target and masking stimuli modulating (nominally) at 0.5 Hz. As can be seen in Figure 8, while the isotropic masking component is attenuated relative to the 1.6 Hz condition (Figure 3), it remains strong. It is worth noting that whereas Meese and Holmes (2007, in press) presented their target/masking stimuli foveally, our study used parafoveal stimulation (0.25–5.25 degrees above and below fixation). Furthermore, while we maintained the spatial and temporal envelopes of target and masking stimuli at all spatial and temporal frequencies tested—thereby confounding the total number of carrier wavelengths present at different spatial and temporal frequencies, both of the other studies scaled envelope size with wavelength. Future research is required to determine whether the distinctive temporal frequency dependencies observed in these studies is due to either stimulus eccentricity, visual angle and/or stimulus duration.

**Functional implications**

The psychophysical threshold elevation effects we observe as a consequence of masking may be expressed as a reduction in signal-to-noise ratio associated with a target-relevant response. It is conceivable therefore that masking may result from either a mask-induced reduction in target-relevant signal response (suppression) and/or and increase in target-irrelevant noise. While the standard linear model of within-channel orientation masking assumes that threshold elevations result from a proportional increase in noise (Graham, 1989), most contemporary models of cross-orientation masking assume a divisive inhibitory process (Boynton & Foley, 1999; Meese & Holmes, 2007). This functional distinction between the mechanisms of orientation-specific and non-specific (isotropic) masking may have important implications for the coding of natural scenes. Numerous studies indicate that the amplitude of natural scenes typically decreases with spatiotemporal frequency (approximately 1/frequency) (Billock & Harding, 1996; Burton & Moorhead, 1987; Field, 1987; Tolhurst, Tadmor, & Chao, 1992). If the low-frequency-biased isotropic masking effect we observe is in fact the result of inhibitory interactions, then arguably this spatial frequency-biased isotropic masking effect may serve to equalize (or “whiten”) the visual system’s response to natural scenes by attenuating the dominant low spatial frequency input. This would have the effect of reducing input redundancy and hence conserving computational and metabolic efficiency (Barlow, 2001; Field, 1987; Srinivasan, Laughlin, & Dubs, 1982). An analogous interpretation was recently applied to the observation that peak amplitude of masking is, on average, smaller at oblique target orientations (45°, 135°) compared to the masking of vertical and horizontal targets (0°, 90°) (Essock, Haun, & Kim, 2009; but see Phillips & Wilson, 1984), thereby perceptually equalizing neural response to the proportionally smaller amplitudes present at oblique orientations in natural scenes.
We acknowledge, however, that the low spatial-frequency-biased isotropic masking effect we observe could conceivably be a consequence of low-spatial-frequency-biased noise (rather than suppression). If this were the case, then we might predict an increased low-frequency response bias (albeit noisy), effectively increasing redundancies evident in the spatial correlational structure of natural scenes. Indeed, an analogous counterinterpretation could be applied to the oblique masking effects reported by Essock et al. (2009).

While this issue remains unresolved, it is interesting to note that orthogonal adaptation has been found to produce greatest threshold elevation in the low spatial (and high temporal) corner of stimulus frequency space (Kelly & Burbeck, 1987). Assuming that cross-orientation adaptation and isotropic (cross-orientation) masking effects are mediated by a common mechanism, the finding that orthogonal adaptation induces profound threshold elevation appears inconsistent with direct cross-orientation inhibition, suggesting a possible role for intervening, possibly disinhibitory, mechanisms. Future research is required to determine the relationship(s) between the mechanisms of masking and adaptation.

**Physiological implications**

That we observe complete interocular transfer of all monoptically derived free parameters (orientation bandwidth, orientation-tuned and untuned masking amplitude) for each spatiotemporal frequency tested might suggest that the neural loci of these masking effects are mediated at a level of processing that receives input from both eyes. It is well established that in primates, the earliest level of visual processing hierarchy that receives substantial binocular input is V1 (Hubel & Wiesel, 1962, 1968). While weak binocular interactions have been observed subcortically in parageniculate and lateral geniculate nuclei of cat (Sengpiel & Vorobyov, 2005; Xue, Carney, Ramoa, & Freeman, 1988), both the strength and completeness of interocular transfer of all free parameters suggests a cortically mediated locus for isotropic and orientation-tuned masking alike.

Primary visual cortex has been found to possess neurons with receptive field properties capable of supporting both orientation-tuned and untuned masking effects. In the case of orientation-tuned masking, both simple and complex cells are obvious candidates (Hubel & Wiesel, 1962, 1968; Ringach et al., 2002). By contrast, the cortical mechanisms of isotropic (untuned) masking may be mediated by populations of cells (possibly complex) with little or no orientation preference (Hirsch et al., 2003) and/or non-specific inhibitory interactions between orientation-tuned cells (Heeger, 1992). Yet another possibility, supported by a recent optical imaging study (MacEvoy, Tucker, & Fitzpatrick, 2009), is that cross-orientation masking effects may not necessarily result from inhibitory interactions between cortical neurons (Heeger, 1992; Morrone et al., 1987) or the addition of saturating non-linearities (Freeman et al., 2002; Li, Peterson, Thompson, Duong, & Freeman, 2005; Li, Thompson, Duong, Peterson, & Freeman, 2006). Rather, the apparent suppression in the response to a given stimulus orientation (physiological or psychophysical) caused by the superimposition of additional (masking) orientations may reflect a redistribution of activity across the population of orientation-selective neurons that preserves the population coded representation of both target and mask orientations while maintaining the average level neural activity across the population. If our psychophysical isotropic masking effects do in fact reflect such a redistribution of V1 activity, one would expect to observe similar spatial frequency contingencies in the magnitude of optically imaged cortical cross-orientation suppression to that observed in the current study. One should also consider the possibility that our psychophysical masking may not necessarily reflect (or even depend upon) the outputs of V1 neurons, as they may be mediated by extra-striate and/or cortico-thalamic interactions (Allison et al., 2001; Carandini et al., 1997; Morrone et al., 1987).

In light of the evidence for strong orientation-tuned masking component in all relevant psychophysical studies of the last 40 years, it is perhaps surprising that chromatically defined (L-M) target and masking stimuli have been found to produce strong isotropic masking, without any evidence of orientation tuning (Medina & Mullen, 2009). This suggests that chromatic masking is mediated by chromatically sensitive neurons with little or no orientation tuning and/or by orientation non-specific interactions between neurons. Neurons that are both chromatically sensitive and untuned for orientation have been found in the earliest primary visual cortical layers as well as the cytochrome oxidase-rich “blob” and “stripe” regions of primate V1 and V2, respectively (Lu & Roe, 2008). Future research will be required to determine whether the chromatic and achromatic forms of isotropic masking are mediated at an equivalent level of visual processing.

The notion that all of our monoptic and dichoptic masking effects are mediated cortically is at odds with recent precortical saturation and thalamo-cortical depression models of physiological XOM (Freeman et al., 2002; Li et al., 2005; Li et al., 2006; Priebe & Ferster, 2006). This interpretation may be premature, however. While the magnitudes of our monoptically and dichoptically derived free parameters are indistinguishable, they may nonetheless be a consequence of different sets of mechanisms. For example, the rate of increase in threshold elevation as a function of orthogonal mask contrast has been found to scale with stimulus speed (TF/SF) in the monocular, but not the dichoptic case (Meese & Baker, 2009), suggesting distinct mechanisms. The masking contrasts employed in our study were constrained to a single value (16× threshold). It is conceivable, therefore, that this particular value may strongly favor masking mechanisms that receive equivalent input from each eye or alternatively, may elicit
an equivalent response from otherwise independent contrast response functions. It is also possible that the dichoptic masking effects we observe may be driven by suppressive interactions, common to both monoptic and dichoptic processes, that occur prior to binocular combination (Meese, Challinor, & Summers, 2008), possibly pre cortical in origin. Again, future research is required to determine the relationship between the orientation-tuned and untuned interocular masking effects observed here and the various monocular, dichoptic, and binocular contrast gain control models (Maehara & Goryo, 2005; Meese, Georgeson, & Baker, 2006).

One factor, which may be critical to our orientation masking results, is the relative spatial phase of the target and masking carrier frequency. Baker and Meese (2007) found evidence for both tuned and untuned components (of similar bandwidths and amplitudes to those observed in our experiment) when carrier modulations of dichoptically presented targets and masks were presented 180° out of phase with each other. When presented in-phase (0° phase difference), however, an additional, narrowly tuned summative component was observed (~8° half width at half height). This additional summative orientation-tuned component was not observed in our experiment, possibly due to the fact that the relative spatial phase of target and masking stimuli was randomized from trial to trial, effectively “blurring out” this component. That spatial phase may be critical for differentiating the mechanisms of monoptic and dichoptic masking is suggested by other studies showing that dichoptic masking thresholds are greater than those observed monoptically when the spatial carrier modulations defining (iso-oriented) target and masking stimuli are in phase (Maehara & Goryo, 2005; Meese et al., 2006). We are currently investigating, in detail, the role of spatial phase in determining the relationship (if any) between the magnitude and orientation bandwidths of monoptic and dichoptic masking mechanisms. Based on these earlier studies, we predict that combining our masking paradigm with a systematic manipulation of spatial phase will reveal the presence of this additional narrow summative orientation-tuned component.

Acknowledgments

J. Cass and D. Alais were supported by Discovery Projects awarded by the Australian Research Council (JC: DP0774697; DA: DP0770299). P. Bex was supported by NIH R01 EY018664 and NIH R01 EY 0119281. IH R01 EY018664 and NIH R01 EY 0119281.

Commercial relationships: none.
Corresponding author: John Cass.
Email: johnc@psych.usyd.edu.au.
Address: School of Psychology, University of Sydney, NSW 2006, Australia.

References


