



# Adaptation and gain pool summation: alternative models and masking data

T.S. Meese \*, D.J. Holmes

*Department of Vision Sciences, School of Life and Health Sciences, Neurosciences Research Institute, Aston University, Aston Triangle, Birmingham B4 7ET, UK*

Received 10 April 2001; received in revised form 10 October 2001

---

## Abstract

Foley [J. Opt. Soc. Am. A 11 (1994) 1710] has proposed an influential psychophysical model of masking in which mask components in a contrast gain pool are raised to an exponent before summation and divisive inhibition. We tested this summation rule in experiments in which contrast detection thresholds were measured for a vertical 1 c/deg (or 2 c/deg) sine-wave component in the presence of a 3 c/deg (or 6 c/deg) mask that had either a single component oriented at  $-45^\circ$  or a pair of components oriented at  $\pm 45^\circ$ . Contrary to the predictions of Foley's model 3, we found that for masks of moderate contrast and above, threshold elevation was predicted by linear summation of the mask components in the inhibitory stage of the contrast gain pool. We built this feature into two new models, referred to as the early adaptation model and the hybrid model. In the early adaptation model, contrast adaptation controls a threshold-like nonlinearity on the output of otherwise linear pathways that provide the excitatory and inhibitory inputs to a gain control stage. The hybrid model involves nonlinear and nonadaptable routes to excitatory and inhibitory stages as well as an adaptable linear route. With only six free parameters, both models provide excellent fits to the masking and adaptation data of Foley and Chen [Vision Res. 37 (1997) 2779] but unlike Foley and Chen's model, are able to do so with only one adaptation parameter. However, only the hybrid model is able to capture the features of Foley's (1994) pedestal plus orthogonal fixed mask data. We conclude that (1) linear summation of inhibitory components is a feature of contrast masking, and (2) that the main aftereffect of spatial adaptation on contrast increment thresholds can be assigned to a single site. © 2002 Elsevier Science Ltd. All rights reserved.

*Keywords:* Inhibition; Contrast normalization; Contrast discrimination; Gain control; Cortex

---

## 1. Introduction

In a masking paradigm, the (simultaneous) presentation of a masking pattern with a test pattern interferes with the detectability of the test pattern causing either threshold elevation (masking) or threshold facilitation. When the mask and test pattern are the same, the paradigm is sometimes called contrast discrimination, and the mask is sometimes called a pedestal. In the adaptation paradigm, prolonged exposure of an appropriate adapting pattern typically raises detection thresholds for a test pattern. Both of these paradigms have been used to estimate parameters of early visual mechanisms, though models of both paradigms have continued to be refined

(e.g. Foley, 1994; Georgeson & Harris, 1984; Mather, Verstraten, & Anstis, 1999). Furthermore, as these processes have become better understood it has become appropriate to devise models that can handle both paradigms (e.g. Foley & Chen, 1997).

One widely reported result in the masking literature is that of the dipper function. Here, detection threshold for the test is measured as a function of mask contrast. When the mask and test patterns are sufficiently similar in orientation, spatial frequency (Legge & Foley, 1980), temporal frequency (Boynton & Foley, 1999) and stimulus onset (Georgeson & Georgeson, 1987), facilitation is found at low mask contrasts and masking is found at higher mask contrasts (these regions are sometimes called the dipper and dipper handle, respectively). The first models of this process employed a static sigmoidal nonlinearity on the output of spatially tuned filters (Legge & Foley, 1980; Nachmias & Sansbury, 1974; Phillips & Wilson, 1984; Wilson, 1980; Wilson,

---

\* Corresponding author. Tel.: +44-121-359-3611x5421; fax: +44-121-333-4220.

E-mail address: t.s.meese@aston.ac.uk (T.S. Meese).

McFarlane, & Phillips, 1983). The nonlinearity typically has the general form of:

$$r = c^p / (z^q + c^q), \quad (1)$$

where  $r$  is the output from the nonlinearity and  $c$  is the contrast response of a linear test filter selective for the test component (and pedestal). The constant  $z$  is  $>0$  and is sometimes referred to as a semi-saturation constant.<sup>1</sup> The model parameters  $p$  and  $q$  control the slopes of the accelerating and compressive parts of the nonlinearity; typically  $p > q$  and  $q \approx 2$ . This model fits contrast discrimination data (and some masking data) because in the low contrast accelerating region,  $\Delta r$  increases for constant  $\Delta c$  and so detection improves, whereas in the higher contrast compressive region,  $\Delta r$  decreases for constant  $\Delta c$  and so detection is hindered. Although very successful with restricted data sets, this model fails when the mask components have very different orientations from the test. For example, Foley (1994) found a substantial amount of masking but no facilitation for a mask component whose orientation was orthogonal to that of the test. This change in response characteristic cannot be accommodated by the single pathway model of Eq. (1) which allows only for lateral translation of the masking function (Foley, 1994). The solution to the problem was inspired by contrast gain control models that had been developed to explain stimulus dependent saturation of neurons in the visual cortex (Albrecht & Geisler, 1991; Heeger, 1992). The crucial step was to add additional terms to the denominator of Eq. (1) representing a weighted ( $w_i$ ) contribution to inhibition from mask components that provide no contribution to the numerator. The group of pathways ( $i = 1:n$ ) that lead to inhibition is often referred to as the contrast gain pool,<sup>2</sup> and typically includes a contribution from the pathway that is itself being inhibited (i.e. self-inhibition). Assuming that excitatory stimulation is provided only by the test component plus pedestal (with contrast proportional to  $c_1$ ), the revised model (referred to as Foley's model 3) can be expressed as

$$r = c_1^p / \left( z^q + \sum [w_i c_i^q] \right). \quad (2)$$

An alternative version (Foley's model 2) sums the  $c_i$  terms before raising them to the exponent  $q$ . Models of one of these two forms have been used by Foley and his colleagues to account for a wide variety of masking data

<sup>1</sup> When  $p = q$ , the function saturates, reaching half of its maximum value when  $c = z$ , hence the term, semi-saturation constant. In the equations presented here, the exponents on the numerator and denominator are not constrained to be equal and so the term is anomalous. Nevertheless, we refer to  $z$  and  $s$  (see forward to Eq. (3)) as semi-saturation constants for convenience.

<sup>2</sup> In this paper we adopt this term for convenience and consistency with other work in the area. We do not address the theoretical implications of its meaning.

in both the spatial (Foley, 1994; Foley & Chen, 1997, 1999) and temporal domains (Boynton & Foley, 1999; Foley & Boynton, 1994) and have also been influential elsewhere (e.g. Itti, Koch, & Braun, 2000; Snowden & Hammett, 1998; Watson & Solomon, 1997).

Foley refers to his model as a functional model because it does not deal with multiple spatial filters explicitly but describes the contributions made by stimulus components to the excitatory (numerator) and inhibitory (denominator) terms in Eq. (2). Upon fitting his model to his data, one important outcome was that the orientation tuning of the inhibitory terms was much broader than the excitatory terms. It is natural to think of the former as the broad selectivity of the gain pool and the latter as the much narrower selectivity of the tuned spatial filter of the detecting mechanism. As Foley (1994) pointed out, excitation of the detecting mechanism always acts to facilitate detection (a *within-channel effect*) whereas masking is caused by the inhibitory effects of the gain pool (a *cross-channel effect*, though a contribution from the test filter is also included). A schematic illustration of this model is shown in Fig. 1a.

In an elegant next step, Foley and Chen (1997) performed experiments involving both masking and adaptation. Their strategy involved comparing fits of their model to masking data both pre- and post-adaptation in order to determine the loci of adaptation by observing which model parameters varied with the state of adaptation. Their conditions were as follows. The test stimulus was always vertical and the variable contrast mask could be either vertical or horizontal. The adapter was either a vertical grating, a horizontal grating or a plaid made from the sum of vertical and horizontal gratings. There was also a condition in which no adaptation took place. Their data are replotted in Figs. 6 and 7 and have been averaged across Foley and Chen's two observers (AHS and CCC), for whom the results were similar. As in previous studies (Foley, 1994; Ross & Speed, 1991; Ross, Speed, & Morgan, 1993), and as outlined above, masking was found regardless of the orientation of the mask but facilitation (a dipper region) was found only when the mask had a similar orientation to the test (Fig. 6). The novel result, however, concerned the aftereffect of adaptation. The main effects occurred when the adapter contained a component that was the same as the test component (i.e. the vertical- and plaid-adapt conditions). In these cases, the entire function was elevated when the mask was horizontal (Fig. 7) but only the dipper region was elevated when the mask was vertical (Fig. 6). In their model, Foley and Chen (1997) concluded that adaptation raises the value of (1) the semi-saturation constant  $z$ , and (2) the weight  $w_i$ , of the horizontal mask in the gain pool.

While the fit to the data was very good, the modelling is worth further consideration. First, as Foley and Chen (1997) pointed out, it is puzzling that adaptation should

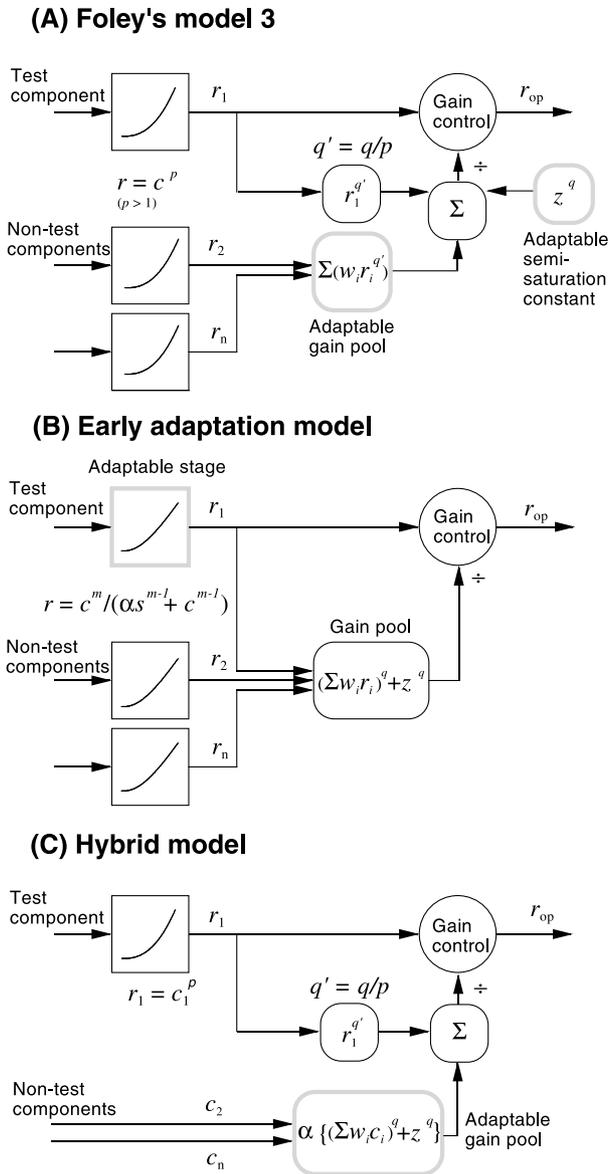


Fig. 1. Three gain control models for masking and adaptation. For all three models, the hashed boxes indicate the loci of adaptation to a stimulus that is processed by the upper pathway. For each hashed box, the state of adaptation is controlled by a single parameter. (A) The type of model used by Foley and Chen (1997). Note the initial non-adaptable expansive nonlinearity, and the two later adaptable stages. (B) Early adaptation model. Note the single stage of adaptation at the threshold-like nonlinearity of the spatial filters. Note also that in (A) the filter nonlinearity is expansive whereas in (B) an initial expansive region is followed by a linear region. (C) Hybrid model. This is very similar to (A) except that: (i) the aftereffects of adaptation are combined into a single stage and (ii) components that are different from the test are summed linearly.

increase the weight of the mask in the gain pool only when the mask is of a different orientation from the test (and the adapter). Second, Foley and Chen (1997) concluded that two parameters were affected by contrast adaptation, leading them to describe a “two-process” model. We wondered whether an alternative model of

their data could be devised in which only a single parameter controlled the state of adaptation. In previous models in which this was so, adaptation was thought to act directly on the outputs of filters stimulated by the adapter (Blakemore & Campbell, 1969; Georgeson, 1985; Georgeson & Harris, 1984; Williams, Wilson, & Cowan, 1982). Examples include the fatigue model, in which the gain of adapted filters is adjusted (e.g. Williams et al., 1982), and a subtractive model, in which a constant is subtracted from the outputs of adapted filters (e.g. Georgeson, 1985). It seemed plausible that a similar arrangement, which we refer to as *early adaptation*, followed by a later nonadaptable gain pool might be able to fit Foley and Chen’s data. Third, we wanted to investigate further the rules for summation in the gain pool. For example, in Foley’s model 3, the terms in the gain pool are raised to an exponent ( $>1$ , and typically  $\geq 2$ ) before summing. A good test of the summing rule is to compare the masking effects of one- and two-component masks that are spectrally remote from the test component and so unlikely to be processed by the same spatial filter as the test. Here, we performed this kind of experiment and found that above a mask contrast of a few percent, summation of mask contrast in the gain pool was linear for a vertical 1 c/deg (or 2 c/deg) test grating and oblique 3 c/deg (or 6 c/deg) mask gratings. We incorporated this feature in two new models: an early adaptation model and a hybrid model of Foley’s model 2 and model 3. Both models contain only a single locus of adaptation (controlled by a single parameter) and provide excellent fits to our data and those of Foley and Chen.

## 2. Experimental methods

### 2.1. Stimuli and observers

Stimuli were displayed using a framestore of a VSG2/3 operating in pseudo-12 bit mode and displayed on a 120 Hz grey-scale monitor (either an Eizo F553-M (mean luminance of 66 cd/m<sup>2</sup>) or Sony Trinitron Multiscan 200PS (mean luminance of 70 cd/m<sup>2</sup>)). Contrast is expressed in dB given by twenty times the log of peak-to-peak stimulus contrast ( $C$ ) given by  $C = 100(L_{max} - L_{min}) / (L_{max} + L_{min})$ . Gamma correction ensured that the monitor was linear over its entire range. A frame interleaving technique was used for test and mask stimuli, meaning that the picture refresh rate was 60 Hz. In Experiment 1, stimulus duration was 100 ms and circular stimuli were matched in size for test and mask with edges curtailed by a raised cosine function. For one spatial condition, stimulus diameter (full width of envelope at half height) was 4.63°, the test component was a 1 c/deg vertical sine-wave grating and the mask was either a plaid with components oriented at  $\pm 45^\circ$  or a

sine-wave grating oriented at  $-45^\circ$ . In a second spatial condition, the stimuli were similar but the test component had a spatial frequency of 2 c/deg, the mask components had a spatial frequency of 6 c/deg and the stimulus diameter was  $2.32^\circ$ . All components were in sine-wave phase with the origin of the display where a small fixation point was located. Experiment 2 was similar to Experiment 1 but a wider range of mask contrasts were used and the temporal and spatial conditions were different. Stimulus duration was 67 ms and both test and mask were phase-reversed mid-way through the presentation (i.e. the temporal envelope was one cycle of a 15 Hz square wave).<sup>3</sup> A 1 c/deg test grating was windowed by a circular Gaussian function with a full-width at half-height of 1.67 cycles of the test stimulus (Foley, 1994). The spatial properties of a 3 c/deg mask were the same as in Experiment 1.

Data were gathered from four observers who had normal or optically corrected-to-normal vision. DJH and TSM are the authors and are both highly practised psychophysical observers. GR and CEB were both unaware of the purpose of the experiment.

## 2.2. Procedure

Contrast detection thresholds for the test grating were measured using a two-interval forced-choice procedure and six interleaved one-up, three-down staircases (Meese, 1995; Wetherill & Levitt, 1965). In a single experimental session, pairs of staircases tracked the threshold for three different conditions: no mask, grating mask and plaid mask, where the peak-to-peak contrast of the plaid mask was matched to that of the grating mask (i.e. the component contrast for the plaid mask was half that of the grating mask). Each staircase terminated after eleven reversals and thresholds were calculated at the 75% correct point on the psychometric functions estimated by probit analysis from the data gathered over the last 10 reversals for each staircase in a pair (i.e. 20 reversals in total). Mask contrast was varied in a pseudo-random order between sessions and thresholds were averaged from five experimental sessions.

## 3. Modelling methods

The model equations were solved numerically using a Pentium PC. We attempted to optimise our fits by using the downhill simplex method (Press, Flannery, Teukolsky, & Vetterling, 1989) initialised with 100 different

pseudo-randomly selected sets of parameters. The reported fits are those that achieved the lowest RMS error, and in all cases were reached from many of the initial parameter sets.

## 4. Experiment 1: linear summation of mask components in the gain pool

### 4.1. Results and discussion

Detection thresholds for a 1 c/deg test grating in the presence of a 3 c/deg grating and plaid are shown in Fig. 2 as a function of the peak-to-peak contrast of the mask. Results for the 2 c/deg grating in the presence of the 6 c/deg masks are shown in Fig. 3. The key feature of all these results is that the functions for the two different

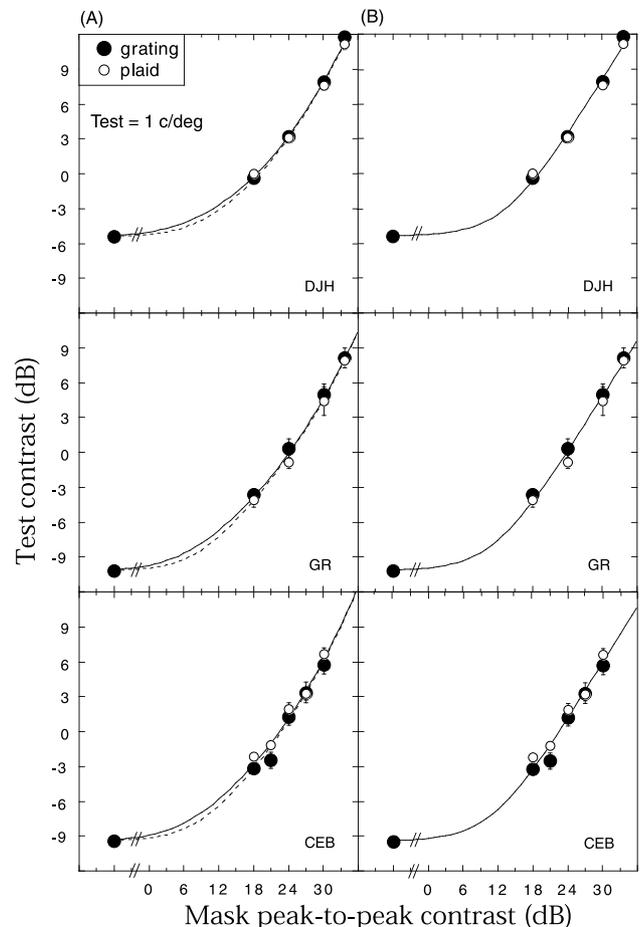


Fig. 2. Masking results for three observers (different rows) and (A) the early adaptation model and (B) the hybrid model. Contrast detection thresholds for a vertical 1 c/deg grating are shown as a function of the peak-to-peak contrast of a 3 c/deg mask that was either a grating (solid symbols) or a plaid (open symbols). Error bars show  $\pm 1$  SE where larger than the symbol. Dashed and solid curves are model fits to the data for grating and plaid masks respectively but are superimposed for the hybrid model. Foley's model 3 predicts that the open symbols should lie well below the filled symbols. See text and Fig. 3 for details.

<sup>3</sup> In this experiment, the phase-reversed test and mask stimulus was used in an attempt to isolate high temporal frequency mechanisms. This was of no theoretical importance in the present paper, but the data form part of a larger body of work for which this manipulation was relevant.

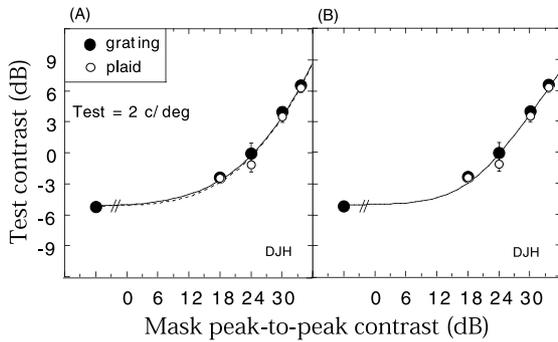


Fig. 3. Similar to Fig. 2, but for a single observer and test and mask spatial frequencies of 2 and 6 c/deg respectively.

masks sit almost exactly on top of one another. This suggests a process equivalent to linear summation of contrast for the two components in the plaid mask. These results are different from others where test and mask spatial frequencies of 3 c/deg (Derrington & Henning, 1989; Holmes & Meese, 1999) and 2 c/deg (Ross & Speed, 1991)<sup>4</sup> were used and where threshold elevation by a plaid mask was found to be greater than that predicted by linear summation of its components. The reasons for this difference remain unclear but one possible complicating factor is the similarity in spatial frequency for mask and test components in the other experiments.

Under the static transducer model of masking (Eq. (1)), one way in which linear summation of mask contrast could occur would be if both mask components provided direct stimulation of the same (linear) filter that responded to the test component. However, this is an unlikely account of our data because the spectral configurations of our stimuli were specifically chosen so that the test and mask components were sufficiently different to be processed by different spatial filters. Furthermore, the static transducer model has been rejected by Foley (1994), who preferred a model incorporating cross-channel inhibition (Eq. (2)). Consequently, a reasonable interpretation of Figs. 2 and 3 is that all or most of the masking is due to the inhibitory effects of the gain pool. However, the form of the gain pool in Foley's model 3 is inappropriate as we show next.

The predicted level of masking (in dB) for the plaid mask relative to the grating mask is shown in Fig. 4 for Foley's model 3 (mathematical details are outlined in Appendix A). Essentially, the two free parameters that are important for these predictions are the exponents  $p$  and  $q$  (see Eq. (2)). In Fig. 4a, the predictions are shown as a function of  $p$  for different values of  $q$ : as  $p$  increases,

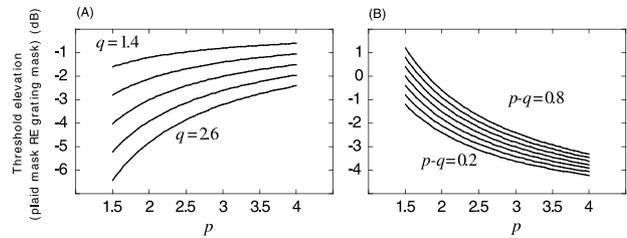


Fig. 4. Threshold elevation for plaid mask relative to grating mask predicted by Foley's model 3. Predictions are shown as a function of  $p$  and different curves are for (A)  $q = 1.4$ – $2.6$  in steps of 0.3 and (B)  $p - q = 0.2$ – $0.8$  in steps of 0.1.

the difference in masking produced by the plaid and the grating decreases. However, in contrast discrimination, for example the exponents  $p$  and  $q$  are subject to the constraint:  $0 < (p - q) < 1$ , where the difference in exponents controls the log–log slope of the upper region of the masking function. Thus, it is helpful to know how the relative levels of masking for plaid and grating depend on  $(p - q)$ . These predictions are shown in Fig. 4b as a function of  $p$ . For a constant difference between  $p$  and  $q$ , the different levels of masking by plaid and grating increases as a function of  $p$ . Published values of  $p$  and  $q$  vary but typically the difference is  $< 0.5$  and  $p$  is  $> 2$  (e.g. Foley, 1994). Fig. 4b shows that, subject to these constraints, the grating mask should produce around 2 dB or more masking than a plaid mask of the same peak-to-peak contrast. There are no pairs of data points in Figs. 2 and 3 close to these predictions (average standard error was  $< 0.5$  dB), causing us to set aside Foley's model 3 and search for an alternative model incorporating linear summation of inhibitory components in the gain pool (see below). Not only is this important for the data reported here but it might also be important for contrast discrimination models intended for broad-band stimuli such as natural images (Tolhurst & Tadmor, 1997). We note, however, that simplifying modelling assumptions, including those concerning the gain pool, have met with some success in this respect (Párraga, Troscianko, & Tolhurst, 2000; Rohaly, Ahumada, & Watson, 1997).

Provision for linear summation is provided in the models of Olzak and Thomas (1999), Ross and Speed (1991) and Foley (1994) model 2. Olzak and Thomas' (1999) model was developed to account for fine spatial discriminations when cues were applied to one or two components. For these tasks, their model predictions were largely unaffected by the summation exponent (see their Appendix A) and so the issue of linear summation was not resolved. Furthermore, as the authors point out, their model saturates with contrast and as such cannot account for the dipper handle in contrast discrimination. Ross and Speed's (1991) model is not as thoroughly tested as other models of similar data (Foley & Chen, 1997) and unlike Foley's model 3, was not sufficiently

<sup>4</sup> Ross and Speed report that the spatial frequency of their test component was 2 c/deg, but report only the orientations ( $\pm 45^\circ$ ) of their grating and plaid masks and results for only one observer.

well developed to account for contrast discrimination in the presence of an additional fixed mask with different orientation from the test (Foley, 1994; Ross et al., 1993). Foley's model 2 has an explicit stage of linear summation of gain pool components and would provide a good fit to our data. However, Foley (1994) reported that this model was inferior to his model 3 for a wide range of masking data, suggesting the need for an alternative model that could accommodate all of these results. We shall return to the limitations of Foley's model 2 later.

## 5. Early adaptation model

In this section we develop the early adaptation model that is fit to our own masking data (Figs. 2 and 3) and the adaptation data of Foley and Chen (1997) (Figs. 6 and 7). The model is not designed to address questions concerning stimulus phase, temporal frequency or probability summation between detecting mechanisms and for ease of comparison with Foley and Chen we adopt Foley's functional approach. Specifically, we express the model in terms of the stimulus components rather than the outputs of multiple filters (Watson & Solomon, 1997). Nevertheless, for all of the data that we model, the mask and test components are either the same or very different, meaning that the pathways in the model (Fig. 1b) can be thought of as representing the groups of filters stimulated by each stimulus component.

Our starting point was to place an adaptable stage at the front end of each pathway in the model consisting of an accelerating nonlinearity followed by an approximately linear region. This addressed the motivation for the work outlined in Section 1 and also allowed for linear summation in the gain pool (at moderate contrast and above), so as to accommodate the results from Experiment 1. This arrangement was also guided by a striking observation in Foley and Chen's data. For a horizontal mask (Fig. 7), the data sets for which there was an adaptation aftereffect were approximately a vertical translation of the unadapted data on double log coordinates. This is equivalent to multiplying the original masking function by some other term. Thus, it seemed likely that we could account for these data with a single adaptation parameter if the locus of adaptation were placed in a separate stage that is multiplied (or divided) by the gain pool. Specifically, the initial adaptable stage of our model is given by

$$r = c^m / (\alpha s^{m-1} + c^{m-1}), \quad (3)$$

where  $s$  is a semi-saturation constant, and  $\alpha$  is a parameter whose magnitude is 1 in the nonadapted state and  $>1$  in the adapted state. Note that Eq. (3) is similar to Eq. (1) but that by arranging that the denominator exponent is one less than the numerator, the function

becomes linear in a region beyond that determined by the semi-saturation constant. This function is shown on double log coordinates in Fig. 5a for  $m = 2.5$ ,  $s = 1$ , and  $\alpha = 1, 2$  and 4. Note that the effect of adaptation is to shift the nonlinear part of the function to the right. This is similar to the effect of a subtractive threshold which has been shown to provide a good account of the adaptation aftereffect on perceived contrast using a contrast matching task (Georgeson, 1985). When the mask

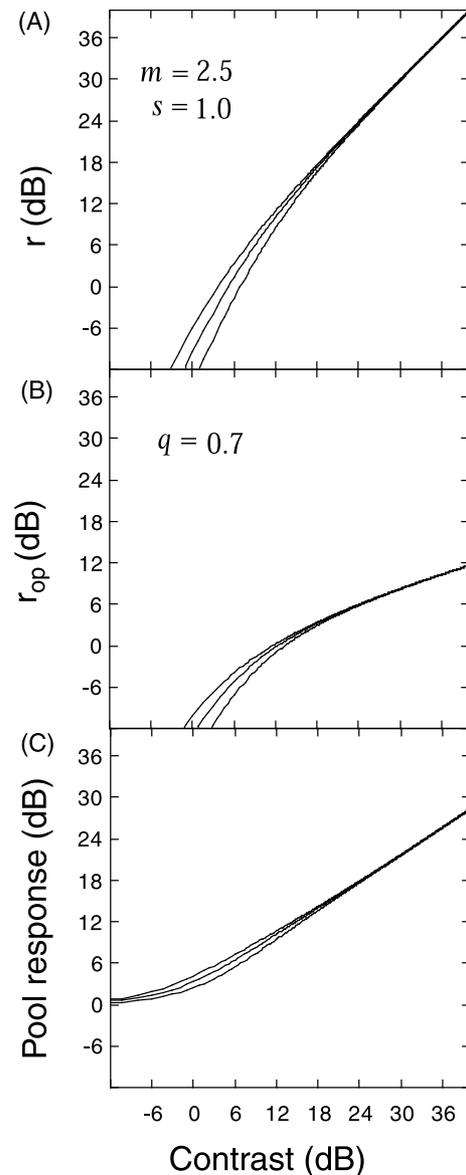


Fig. 5. Responses at different stages of the early adaptation model for three different states of adaptation (curves from left to right are for  $\alpha = 1, 2$ , and 4). (A) Contrast response of the initial adaptable stage as a function of test contrast. (B) Similar to (A) but after the divisive inhibition stage. In this case, the mask component has the same spatial frequency, orientation and phase as the test component (i.e. it is a pedestal). (C) Contrast response of the gain pool (the denominator in Eq. (4)) for three different states of adaptation of the input to the gain pool.

is the same as the test,  $r_1$  is computed and  $c$  is the sum of mask and test contrasts. When the mask is different from the test,  $r_1$  and  $r_{\text{mask}}$  are computed, where  $c$  is the contrast of the test and mask respectively. When the number of mask components is greater than one (as for the plaid mask in our experiments), then  $r_{\text{mask}}$  is computed for each of the  $i = 2:n$  mask components (see Fig. 1b). The adaptation parameter  $\alpha_i \neq 1$ , only if there has been adaptation to the  $i$ th component.

The second stage of the model is divisive inhibition. This involves dividing  $r_1$  by the weighted ( $w_i$ ) linear sum of the responses to each of the components in the stimulus to generate the model output ( $r_{\text{op}}$ ):

$$r_{\text{op}} = r_1 / \left\{ z^q + \left( r_1 + \sum [w_i r_{\text{mask},i}] \right)^q \right\}. \quad (4)$$

A schematic illustration is shown in Fig. 1b.

We assume that a contrast increment in the test is detected when the difference in  $r_{\text{op}}$  computed for the mask and mask + test intervals equals or exceeds some criterion  $k$ . That is, at contrast increment threshold:

$$r_{\text{op mask+test}} - r_{\text{op mask}} = k. \quad (5)$$

By trivially arranging that the semi-saturation constants in Eqs. (3) and (4) are equal ( $s = z$ ) and that  $w_i$  is equal for all  $i$  (e.g. it is the same for each of the plaid mask components), the number of free parameters in the model is four or five depending on whether the mask is the same or different from the test, increasing to five or six when adaptation is modelled (Figs. 6 and 7 and Table 1). This compares favourably with Foley and Chen's model which has two free parameters that depend upon each of their four adaptation conditions (though in some cases the dependencies are only marginally different), plus a further three, giving 11 free parameters in total.

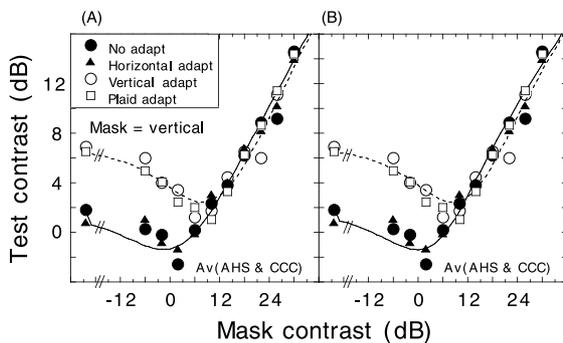


Fig. 6. Vertical mask data averaged from AHS and CCC are replotted from Foley and Chen (1997) and fit by the early adaptation model (A) and the hybrid model (B) described in the text. The fit is the same for the 'no adapt' and the 'horizontal adapt' conditions and is shown by the solid curve. The fit is also the same for the 'vertical adapt' and the 'plaid adapt' conditions and is shown by the dashed curve. The models were fit simultaneously to these data and those shown in Fig. 7.

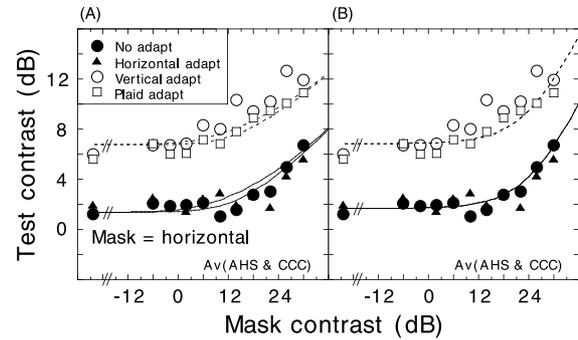


Fig. 7. Horizontal mask data averaged from AHS and CCC are replotted from Foley and Chen (1997) and fit by the early adaptation model (A) and the hybrid model (B) described in the text. The fits for the 'no adapt' and the 'horizontal adapt' conditions are shown by the solid curves and those for the 'vertical adapt' and the 'plaid adapt' conditions are shown by the dashed curve. In (A), the curve pairs are slightly misaligned with the 'no adapt' and 'vertical adapt' curves being the higher. The models were fit simultaneously to these data and those shown in Fig. 6.

### 5.1. Model behaviour and model fits

With the same parameters as used in Fig. 5a, the response of the model is shown in Fig. 5b as a function of test contrast for  $q = 0.7$  and for a mask matched to the test. The model output is very similar to that which is achieved using Eqs. (1) and (2), and not surprisingly, it provides a good account of conventional contrast discrimination data (filled circles and solid curve in Fig. 6a). Note also that the main effect of adapting to a (high contrast) test component is to reduce sensitivity when activity in the excitatory pathway is low, but to leave the high contrast limb of the response function intact (Fig. 5b). Consequently, the model does a good job in modelling the adaptation aftereffect when the test is matched to the mask (open symbols and dashed curve in Fig. 6a).

When the mask and test are different, the masking effect is due to the inhibitory action of the gain pool (the denominator in Eq. (4)) which is shown as a function of mask contrast in Fig. 5c with  $w = 1$ . There is no facilitation because this requires  $r_1$  in the numerator of Eq. (4) to be driven up the accelerating part of its function. Also, masking becomes effective only at moderate contrasts and above because of the semi-saturation constant  $z$  and the functions can be shifted laterally by changing the inhibitory weight  $w$ . These features mean that the model captures the effect of an orthogonal mask very well (solid circles and lower solid curve in Fig. 7a). The small effect of adapting to the orthogonal mask orientation is evident from Fig. 5c and has its greatest impact at low mask contrast. Consequently, when the mask and test are different, adapting to the mask will have little or no effect on masking as shown in Fig. 7a by the small differences between the pairs of dashed and solid curves. Thus, from these data alone, there is no need to suppose

Table 1  
Early adaptation model parameters and goodness of fit

	$k$	$m$	$q$	$z = s$	$w$	$\alpha$	RMS error (dB)	Number of data points	Linear slope for grating mask	Linear slope for plaid mask	Foley (1994) model 3 RMS error (dB)
DJH (Fig. 2)	0.14	2.5 (fixed)	1.19	1 (fixed)	0.22	–	0.25	9	0.79	0.73	0.942
GR (Fig. 2)	0.045	2.5 (fixed)	1.29	1 (fixed)	0.35	–	0.26	9	0.72	0.73	0.755
CEB (Fig. 2)	0.055	2.5 (fixed)	1.30	1 (fixed)	0.35	–	0.62	11	0.71	0.66	1.236
DJH (Fig. 3)	0.15	2.5 (fixed)	1.29	1 (fixed)	0.09	–	0.32	9	0.57	0.57	0.65
DJH (Fig. 8)	0.33	2.5 (fixed)	1.05	1 (fixed)	0.43	–	0.52	13	0.92	0.88	1.136
TSM (Fig. 8)	0.34	2.5 (fixed)	0.98	1 (fixed)	0.46	–	0.54	19	0.78	0.77	0.694
AHS and CCC (Figs. 6 and 7)	0.24	2.24	0.61	1.61	0.22	5.32	0.91	88	–	–	–
AHS and JYS (Fig. 9)	0.26	3.05	0.79	3.71	0.92	–	1.82	22	–	–	–

Slopes for best fitting straight lines through the high contrast ( $\geq 18$  dB) part of the masking functions and goodness of fit for Foley's model 3 are also shown. Results are for the four observers in Experiments 1 and 2, the average of AHS and CCC replotted from Foley and Chen (1997) and the average of AHS and JYS replotted from Foley (1994).

that the input to the gain pool does not pass through the primary site of adaptation (Foley & Chen, 1997).

When the mask is different from the test and the adaptor matches the test, unlike in Fig. 6, the adaptation aftereffect is not abolished at high mask contrasts. This is because in this case, there is no pedestal to raise activity in the test pathway away from the aftereffect of adaptation (Fig. 5b). Consequently, the model provides a good account of the data for these conditions (Fig. 7a, dashed curves).

The early adaptation model's parameters and goodness of fit are shown in Table 1 (a full summary of all symbols used in the main body of this paper is provided in Appendix B). The goodness of fit achieved by our own model (RMS error = 0.91 dB) is comparable with that estimated from Foley and Chen by combining the separate error measures for their two observers (RMS error = 1.03 dB). The small difference between these figures could be due to the data averaging across observers, small errors in the replotted data or slight differences in the efficacy of the models. The main point, however, is that the early adaptation model achieved a similar goodness of fit to the model of Foley and Chen (1997) but with fewer parameters. In particular, only a single parameter for controlling the state of adaptation was needed.

The early adaptation model was also fit to the grating and plaid masking data from Experiment 1 with two of the five free parameters fixed ( $m = 2.5$ ,  $z = 1$ ). Not surprisingly, the approximately linear summation of mask contrast in the gain pool meant that the fit of the model was very good (Table 1).

## 6. Experiment 2: nonlinear summation for low contrast masks

### 6.1. Results and discussion

Because of the nonlinearity feeding into the input of the gain pool in the early adaptation model (Fig. 1a), the model predicts that a grating should produce slightly more masking than a plaid at low mask contrasts. We tested this prediction by repeating the experiment over a wider range of mask contrasts.

Results are shown in Fig. 8. Over the high contrast limb of the functions, masking by plaid and grating is very similar, but around a contrast of 12 dB (4%), there is slightly more masking by the grating than by the plaid. This was confirmed by two-tailed matched *t*-tests applied to the plaid and grating data at each mask contrast. (Single and double asterisks denote reliable differences at significance levels of  $p < 0.05$  and  $p < 0.01$ , respectively.) These differences are predicted by the early adaptation model, for which parameter values and goodness of fit are shown in Table 1.

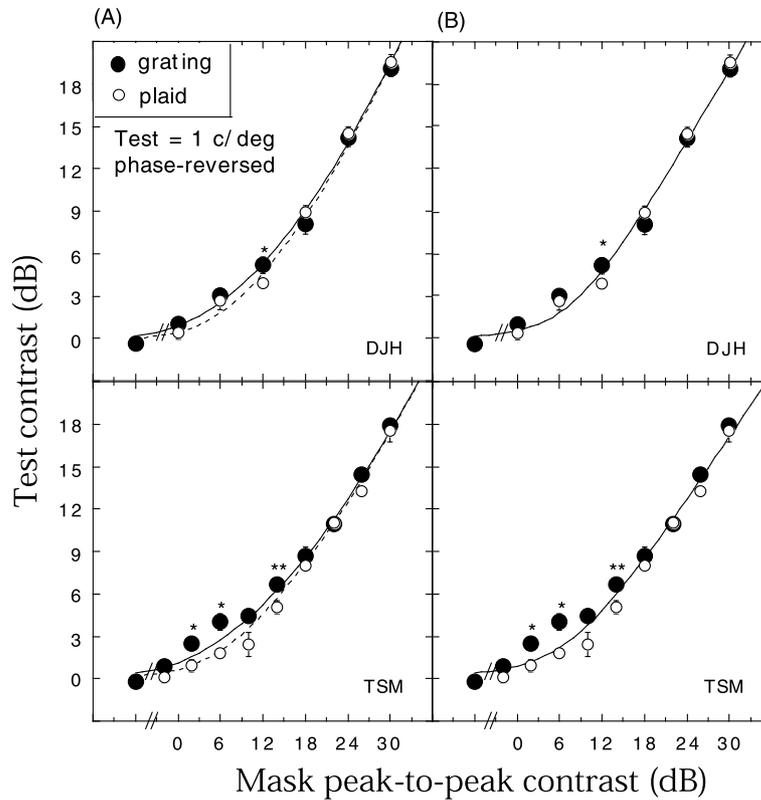


Fig. 8. Similar to Fig. 2, but for a wider range of mask contrasts and for mask and test stimuli that were a single cycle of a 15 Hz square wave. Two-tailed matched *t*-tests were performed on all of the data pairs (open and closed symbols). An asterisk and double asterisk denote significant differences at  $p < 0.05$  and  $p < 0.01$  respectively. In (A) the data are fit by the early adaptation model and in (B) they are fit by the hybrid model. See text for details.

### 7. Problems with the early adaptation model

Table 1 shows that in Experiments 1 and 2, better fits were achieved by the early adaptation model than by Foley’s model 3 (Eq. (2)), which in these cases had the same number of free parameters. (The three free parameters were  $q$ ,  $w_2$ , and  $k$  (Eq. (5)) and the fixed parameters were  $p = 2.5$ ,  $z = 1$ , and  $w_1 = 1$ ). Note that for the early adaptation model  $q$  tended to be greater than one. In conventional contrast discrimination experiments in which the mask and test are processed by the same pathway, the log–log slope of the upper region of the masking function is usually found to be less than one (e.g. see Fig. 9). This part of the function is well approximated by  $1 - (p' - q)$ , where  $p'$  is the effective exponent of the excitatory term and  $q$  is the exponent of the inhibitory term. In the early adaptation model  $p'$  is approximately one because the excitatory term is large and the semi-saturation constant in Eq. (3) has negligible effect. Therefore,  $q$  is equal to the log–log slope of the data, and is less than one. Thus, for the early adaptation model, a single value of  $q$  cannot fit the masking functions reported here as well as conventional contrast discrimination functions. We suggest three possible solutions to this conflict. First, if the pathway

for the test component is made slightly sensitive to the remote mask components (Ross & Speed, 1991), then this pushes the excitatory term into its linear range of Eq. (3) and the log–log slope is given by  $q$ . Second, the model could be modified to allow for different values of  $q$  from different spectral regions of the contrast gain pool. Third,  $q$  could be made dependent upon  $p'$ .

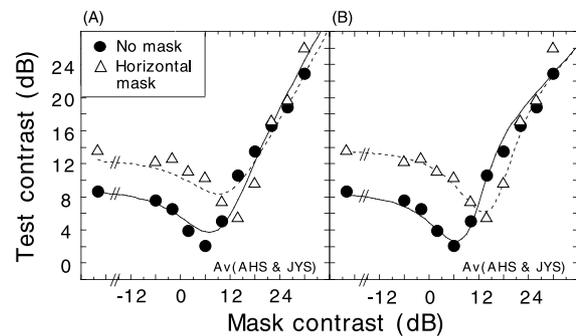


Fig. 9. Data averaged across AHS and JYS from Foley (1994) and fit by the early adaptation model (A) and the hybrid model (B) described in the text. The figures show contrast increment thresholds (test contrast) in the absence of any other masker (solid circles) and in the presence of a horizontal mask grating with contrast of 10% (open triangles). The fits for the ‘no mask’ and ‘horizontal mask’ are fit by the solid curves and dashed curves respectively.

There is, however, another problem with the early adaptation model. Foley (1994) investigated the effects of a fixed contrast (10%) orthogonal mask on contrast discrimination. He found the fixed mask had little or no effect on the dipper handle but that the region of facilitation was shifted both vertically and horizontally. Consequently, there was a small region where the fixed mask actually improved contrast discrimination. This is a distinct feature of Foley's data (which have been replotted in Fig. 9) and Foley's model 3, though we note it is barely noticeable in the data of Ross et al. (1993) who performed a similar experiment. In our own laboratory we have found this feature to be pronounced for one observer, but not for a second (Holmes & Meese, 2001), suggesting that the discrepancy between Foley (1994) and Ross et al. (1993) could be due to observer differences. Nevertheless, for neither the early adaptation model nor Foley's model 2 were we able to find parameter sets that captured the presence of this feature. This is shown in Fig. 9a for the early adaptation model (see Table 1 for parameter values), which produced an RMS error of 1.82 dB and is slightly better than our best fit of Foley's model 2 which achieved an RMS error of 1.98 dB (not shown). Foley's model 3 provides a good fit to these data (see Foley, 1994), but as this model mischaracterises our own data, we sought an alternative solution.

## 8. Alternative models

### 8.1. Hybrid model

One alternative model that does not suffer the problems of the early adaptation model is that shown in Fig. 1c. Like Foley's model 3, the self-inhibiting component is raised to an exponent ( $q$ ) before being combined with the other gain pool components. However, like Foley's model 2, the remaining components in the gain pool are linearly summed before being raised to the exponent  $q$ . Formally, the model is given by

$$r = c_1^p / \left\{ \alpha \left( z^q + \left[ \sum \{w_i c_i\} \right]^q \right) + c_1^q \right\}, \quad (6)$$

where  $c_1$  is the contrast of the test component (plus pedestal) and  $c_i$  is the contrast of the  $i$ th mask component ( $i = 2:(n+1)$ ) that is different from the test.

The model provides a good fit to the grating and plaid masking data (Figs. 2b, 3b and Table 2). For the conditions in Foley (1994) fixed mask experiment, the hybrid model is identical to Foley's model 3 and, as found by Foley (1994), fits those data well (Fig. 9b and Table 2). The hybrid model readily lends itself to the inclusion of just a single adaptation parameter ( $\alpha$  in Eq. (6)) which serves as a coefficient to the two parameters identified by Foley and Chen as being subject to an adaptation aftereffect ( $z$  and  $w$ ). The hybrid model's fit to Foley and Chen's masking and adaptation data is shown in Figs. 6b and 7b and Table 2 and is good. One minor weakness of the hybrid model is that the late stage of adaptation means that the model predictions are exactly the same for the no adapt and the horizontal adapt conditions (open symbols, Fig. 5) and the vertical adapt and plaid adapt conditions (filled symbols, Fig. 5). The differences in the data for these conditions are small, but were noted by Foley and Chen (1997).

One interpretation of the hybrid model is that the linear pooling of mask components takes place within a single linear filter (prior to an output nonlinearity), that is sensitive to both components of the mask. Visual neurons in the retina and LGN have circular receptive fields and could, in principle, perform the necessary summation. However, there is one minor quantitative weakness of the hybrid model that might cast some doubt on this interpretation. Because summation in the gain pool is linear, the hybrid model does not capture the slight differences in masking by a grating and a plaid at low mask contrasts (Fig. 8b). As already discussed in Experiment 2, one interpretation of these data is that, prior to summation, each mask component passes through a nonlinearity that has its greatest impact at low contrast (as in the early adaptation model). Such a feature is inconsistent with early summation of mask contrast in circular filters. On the other hand, we note that our plaid masks contained fewer regions of maximum contrast than our corresponding grating masks.

Table 2  
Hybrid model parameters and goodness of fit for the same data as in Table 1

	$k$	$p$	$q$	$z$	$w$	$\alpha$	RMS error (dB)	Number of data points
DJH (Fig. 2)	0.17	2.5 (fixed)	2.06	1 (fixed)	0.22	–	0.43	9
GR (Fig. 2)	0.05	2.5 (fixed)	2.03	1 (fixed)	0.27	–	0.42	9
CEB (Fig. 2)	0.06	2.5 (fixed)	2.04	1 (fixed)	0.28	–	0.53	11
DJH (Fig. 3)	0.17	2.5 (fixed)	1.93	1 (fixed)	0.12	–	0.43	9
DJH (Fig. 8)	0.51	2.5 (fixed)	2.21	1 (fixed)	0.44	–	0.63	13
TSM (Fig. 8)	0.52	2.5 (fixed)	1.97	1 (fixed)	0.49	–	0.77	19
AHS and CCC (Figs. 6 and 7)	0.41	1.99	1.50	1.73	0.10	3.75	0.92	88
AHS and JYS (Fig. 9)	0.67	3.47	3.05	2.9	0.57	–	1.17	22

This might be influential in the masking process, though why only at low contrasts is not clear.

## 8.2. Fatigue model

Finally, we return to adaptation and consider a widely used version of an adaptation model known as the fatigue model. The version considered here is another form of early adaptation. In this case, the contrast term  $c_1$  on both the numerator and denominator of Eq. (2) is attenuated, equivalent to a multiplicative reduction in effective contrast in the adapted pathway. Similarly, adaptation to other pathways in the gain pool results in multiplicative reduction of their inhibitory effects. Foley and Chen (1997) and others have rejected this model but for completeness, we discuss its failings here. The model predicts that in contrast discrimination (Fig. 6), adaptation to the test pathway moves the dipper-function obliquely on log–log coordinates because of the multiplicative loss in effective contrast (e.g. see Wilson & Humanski, 1993). To a first approximation this is a fair account of the data in Fig. 6, though it fails in one important respect. Because the dipper handle is well fit by a straight line on log–log axes with a slope less than one (e.g. the replotted dipper handles in Fig. 6 have slopes between 0.59 and 0.73), the fatigue model predicts that adaptation must raise threshold elevation in this region of the function by a constant log increment. However, this is not consistent with experimental data. Some studies have found pre- and post-adaptation dipper handles to be superimposed (Foley & Chen, 1997; Määttänen & Koenderink, 1991; Ross & Speed, 1991; Ross et al., 1993) whereas others have found differences in their slopes (Greenlee & Heitger, 1988; Wilson & Humanski, 1993). In all six of these studies, regions were found in which contrast discrimination was actually improved by adaptation. Taken together, these qualitative failings cast doubt on the fatigue model, though quantitatively, the model fares reasonably well with the data replotted here. Our best fit of the fatigue model to the data in Fig. 6 produced an RMS error of 1.12 dB, which was a little worse than the RMS error of 0.85 dB achieved by the early adaptation model fit to the same data. However, unlike the early adaptation model (and the hybrid model), the two curves did not cross, the regions of facilitation were shallow and in the region of the dipper handle, the curves straddled the data, being separated by  $\approx 1.3$  dB.

## 9. Summary and conclusions

We have found evidence for linear summation of mask components prior to divisive inhibition when the mask components have different spatial frequency and orientations from the test. This result is inconsistent

with Foley's model 3, but is a feature of two new models of masking—an early adaptation model and a hybrid model based on Foley's models 2 and 3. Both models provide a good account of Foley and Chen's masking and adaptation data and the data presented here, though the early adaptation model provides a better qualitative account of some of the subtleties in the data. On the other hand, only the hybrid model is able to capture the features of Foley (1994) fixed mask data.

Both new models have only a single parameter to control the state of adaptation. In the early adaptation model, this parameter operates directly on the output of adapted linear filters, whereas in the hybrid model, it acts on all of the terms in the gain pool other than the contribution from the adapted pathway. This second model extends the work of Foley and his colleagues in several important ways. First, it recognises that the two adaptable terms identified by Foley and Chen need not be thought of as two separate processes, but can be combined into a single adaptable stage. Second, the previous observation that adaptation does not attenuate self-inhibition (Foley & Chen, 1997) suggests architecture consistent with the view that the self-inhibitory component is also treated differently in the summing stage of the gain pool.

## Acknowledgements

We would like to thank Steve Anderson and Mark Georgeson for helpful comments and suggestions. We would also like to thank Mark Georgeson for writing the software used in replottting the digitally scanned data of Foley (1994) and Foley and Chen (1997).

## Appendix A

Here we develop the predictions made by Foley (1994) model 3 for Experiments 1 and 2. Foley's model 3 can be expressed by Eq. (2) in the main body of the paper. At moderate mask contrasts and above, the mask components dominate the inhibitory terms and so the contribution from  $z$  can be ignored. Thus, the model predicts that

$$(c_{t\_plaid}^p)/(2w[c_{mask}/2]^q) = (c_{t\_grat}^p)/(wc_{mask}^q)$$

where  $c_{t\_plaid}$  and  $c_{t\_grat}$  are the contrast detection thresholds for the test component in the presence of the grating and plaid masks respectively and  $c_{mask}$  is the contrast of the mask. This rearranges to give

$$c_{t\_plaid}/c_{t\_grat} = (2[1/2]^q)^{1/p}$$

and is the ratio plotted in Fig. 4 where it is expressed in dB.

## Appendix B

Summary of symbols used in the main body of the paper.

$r_{op}$	model response, after gain control.
$r_i$	response of the $i$ th pathway after early static nonlinearity.
$r_{mask\_i}$	as above, but with the qualification that the stimulus is the $i$ th mask component.
$c_i$	linear contrast response of the $i$ th pathway, prior to any nonlinearities and equivalent to component contrast. When test and mask components are the same this is the sum of their contrasts.
$z, s$	semi-saturation constants.
$p, p', m$	numerator exponents.
$q, q'$	denominator exponents.
$w_i$	weight of the $i$ th component in the gain pool.
$\alpha$	adaptation parameter. ( $\alpha = 1$ in unadapted state and $\alpha > 1$ in adapted state.)
$k$	model response difference between mask stimulus and test plus mask stimulus required to achieve detection.
$n$	number of pathways contributing to the gain pool.
$C$	Michelson contrast of sine-wave components used in the experiments.

## References

- Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity and the contrast response function of simple cells in the visual cortex. *Visual Neuroscience*, 7, 531–546.
- Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, 203, 237–260.
- Boynton, G. M., & Foley, J. M. (1999). Temporal sensitivity of human luminance pattern mechanisms determined by masking with temporally modulated stimuli. *Vision Research*, 39, 1641–1656.
- Derrington, A. M., & Henning, G. B. (1989). Some observations on the masking effects of two-dimensional stimuli. *Vision Research*, 29, 241–246.
- Foley, J. M. (1994). Human luminance pattern vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A*, 11, 1710–1719.
- Foley, J. M., & Boynton, G. M. (1994). A new model of human luminance pattern vision mechanisms: analysis of the effects of pattern orientation, spatial phase, and temporal frequency. In T. A. Lawton (Ed.), *Computational Vision based on Neurobiology*, SPIE Proceedings 2054 (pp. 32–42).
- Foley, J. M., & Chen, C.-C. (1997). Pattern detection in the presence of maskers that differ in spatial phase and temporal offset: threshold measurements and a model. *Vision Research*, 37, 2779–2788.
- Foley, J. M., & Chen, C.-C. (1999). Pattern detection in the presence of maskers that differ in spatial phase and temporal offset: threshold measurements and a model. *Vision Research*, 39, 3855–3872.
- Georgeson, M. A. (1985). The effect of spatial adaptation on perceived contrast. *Spatial Vision*, 1, 103–112.
- Georgeson, M. A., & Georgeson, J. M. (1987). Facilitation and masking of briefly presented gratings: time course and contrast dependence. *Vision Research*, 27, 369–379.
- Georgeson, M. A., & Harris, M. G. (1984). Spatial selectivity of contrast adaptation: models and data. *Vision Research*, 24, 729–741.
- Greenlee, M. W., & Heitger, F. (1988). The functional role of contrast adaptation. *Vision Research*, 28, 791–797.
- Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9, 181–197.
- Holmes, D. J., & Meese, T. S. (1999). Simultaneous masking by plaids and gratings: threshold elevation and threshold facilitation. ECVF Abstract. *Perception*, 28 (suppl.), 88.
- Holmes, D. J., & Meese, T. S. (2001). Linear summation for remote masks in a contrast gain pool. ECVF Abstract. *Perception*, 30 (suppl.), 81.
- Itti, L., Koch, C., & Braun, J. (2000). Revisiting spatial vision: toward a unifying model. *Journal of the Optical Society of America A*, 17, 1899–1917.
- Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, 70, 1458–1471.
- Määttänen, L. M., & Koenderink, J. J. (1991). Contrast adaptation and contrast gain control. *Experimental Brain Research*, 87, 205–212.
- Mather, G., Verstraten, F., & Anstis, A. (1999). *The Motion Aftereffect: A Modern Perspective*. London: MIT Press.
- Meese, T. S. (1995). Using the standard staircase to measure the point of subjective equality: a guide based on computer simulations. *Perception and Psychophysics*, 57, 267–281.
- Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: discrimination may be better than detection. *Vision Research*, 14, 1039–1041.
- Olzak, L. A., & Thomas, J. P. (1999). Neural recoding in human pattern vision: model and mechanisms. *Vision Research*, 39, 231–256.
- Párraga, C. A., Troscianko, T., & Tolhurst, D. J. (2000). The human visual system is optimised for processing the spatial information in natural visual images. *Current Biology*, 10, 35–38.
- Phillips, G. C., & Wilson, H. R. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *Journal of the Optical Society of America*, 1, 226–232.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1989). *Numerical Recipes in Pascal: The Art of Scientific Computing*. Cambridge: Cambridge University Press.
- Rohaly, A. M., Ahumada, A. J., & Watson, A. B. (1997). Object detection in natural backgrounds predicted by discrimination performance and models. *Vision Research*, 37, 3225–3235.
- Ross, J., & Speed, H. D. (1991). Contrast adaptation and contrast masking in human vision. *Proceedings of the Royal Society of London B*, 246, 61–69.
- Ross, J., Speed, H. D., & Morgan, M. J. (1993). The effects of adaptation and masking on incremental thresholds for contrast. *Vision Research*, 33, 2051–2056.
- Snowden, R. J., & Hammett, S. T. (1998). The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. *Vision Research*, 38, 1935–1945.
- Tolhurst, D. J., & Tadmor, Y. (1997). Discrimination of changes in the slopes of the amplitude spectra of natural images: band-limited contrast and psychometric functions. *Perception*, 26, 1011–1025.
- Watson, A. B., & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America*, 14, 2379–2391.
- Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, 18, 1–10.

Williams, D. W., Wilson, H. R., & Cowan, J. D. (1982). Localized effects of spatial-frequency adaptation. *Journal of the Optical Society of America*, 72, 878–887.

Wilson, H. R. (1980). A transducer function for threshold and suprathreshold human vision. *Biological Cybernetics*, 38, 171–178.

Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, 23, 873–882.

Wilson, H. R., & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vision Research*, 33, 1133–1149.