Georgeson, Wallis, Meese & Baker (2016) *Vision Research*
doi: 10.1016/j.visres.2016.08.001

Contrast and Lustre: a model that accounts for eleven different forms of contrast discrimination in binocular vision

Mark A. Georgeson, Stuart A. Wallis, Tim S. Meese, Daniel H. Baker**

School of Life & Health Sciences, Aston University, Birmingham, UK
**Department of Psychology, University of York, Heslington, York, YO10 5DD, UK
Correspondence email: m.a.georgeson@aston.ac.uk

Abstract

Our goal here is a more complete understanding of how information about luminance contrast is encoded and used by the binocular visual system. In two-interval forced-choice experiments we assessed observers' ability to discriminate changes in contrast that could be an increase or decrease of contrast in one or both eyes, or an increase in one eye coupled with a decrease in the other (termed *IncDec*). The base or pedestal contrasts were either in-phase or out-of-phase in the two eyes. The opposed changes in the *IncDec* condition did not cancel each other out, implying that along with binocular summation, information is also available from mechanisms that do *not* sum the two eyes' inputs. These might be monocular mechanisms. With a *binocular* pedestal, monocular increments of contrast were much easier to see than monocular decrements. These findings suggest that there are separate binocular (*B*) and monocular (*L,R*) channels, but only the largest of the three responses, *max*(*L,B,R*), is available to perception and decision. Results from contrast discrimination and contrast matching tasks were described very accurately by this model. Stimuli, data, and model responses can all be visualized in a common *binocular contrast space*, allowing a more direct comparison between models and data. Some results with out-of-phase pedestals were *not* accounted for by the *max* model of contrast coding, but were well explained by an extended model in which gratings of opposite polarity create the sensation of *lustre*. Observers can discriminate changes in lustre alongside changes in contrast.

Keywords: contrast discrimination, binocular vision, dichoptic masking, contrast matching, binocular lustre, computational model

1 Introduction

1.1 Functional architecture of binocular vision from psychophysics

Two eyes are better than one, but not always. Observers with normal binocular vision typically show faster reaction times, better spatial acuity and higher contrast sensitivity using two eyes rather than one (for reviews see Blake, Sloane, & Fox, 1981; Blake & Fox, 1973). When measured with forced-choice techniques, contrast thresholds with one eye are on average 1.6 to 1.7 times higher than with two eyes (Meese, Georgeson, & Baker, 2006; Simmons & Kingdom, 1998; Simmons, 2005) - consistently higher than the classical figure of √2 (1.41) (Campbell & Green, 1965). It seems clear that this *binocular advantage* in visual performance arises from *binocular summation* of signals from each eye (Fig. 1a), carried out by binocular cells in the primary visual cortex (Hubel & Wiesel, 1962; Anzai, Bearse, Freeman, & Cai, 1995).

Surprisingly however, the binocular advantage in a variety of spatial tasks (Landolt C acuity, letter recognition, orientation discrimination) tends to evaporate at higher contrasts (Bearse & Freeman, 1994; Home, 1978). We focus here on another simple visual task – *contrast discrimination* – which also appears to show no binocular advantage. The task is to decide which of two otherwise-identical sinewave gratings has the higher contrast. When the base or *pedestal* contrast (C) is above threshold, then the contrast difference ΔC required to distinguish the two contrasts, C and C+ΔC, is the same whether the test gratings are shown to one eye or to both eyes (Legge, 1984; Maehara & Goryo, 2005; Meese, Georgeson, & Baker, 2006). This may seem paradoxical, but it does *not* imply that

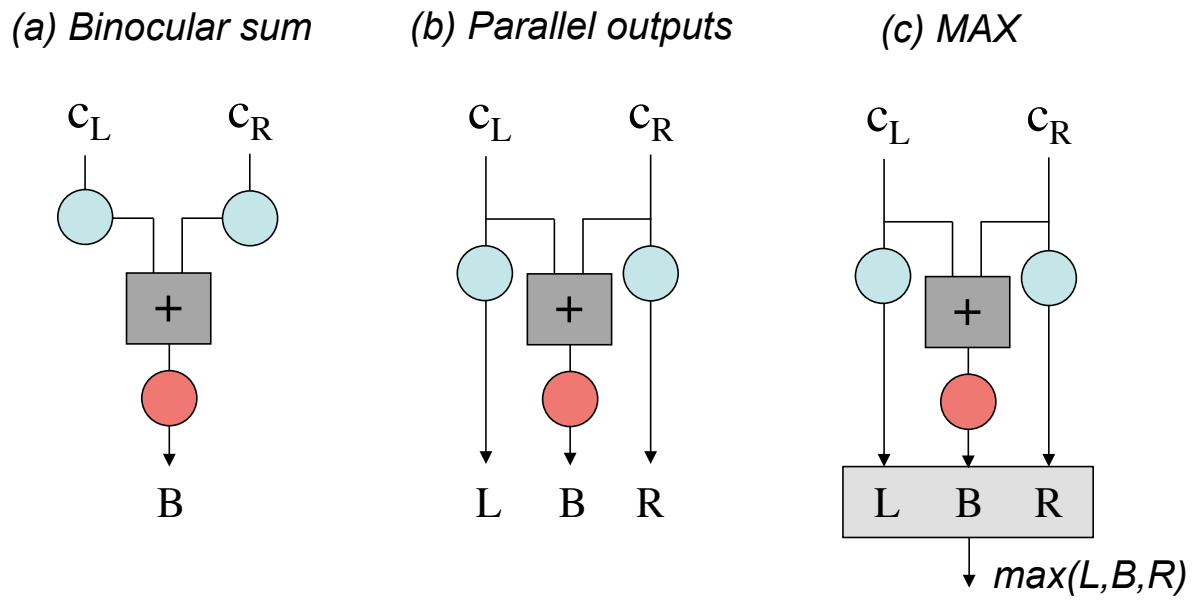

Figure 1. Some basic ideas about binocular combination. (a) Binocular summation: a single binocular output channel (B, red) combines monocular responses to contrasts (c_L, c_R) in the left and right eyes. Blue disks are monocular units. (b) Monocular outputs (L,R) in parallel with the binocular one. (c) In this paper we explore the idea that parallel outputs are available initially, but only the largest of them, *max(L,B,R),* is selected for further processing.

binocular summation is absent above threshold. Rather, this and related results reveal that the process of binocular summation is accompanied by a process of *interocular suppression* that operates in addition to the *self-suppression* that is common in contrast gain control models of contrast discrimination (e.g. Legge & Foley, 1980). When the same image is in both eyes, the benefit of binocular summation is almost exactly offset by the doubling of suppression, leaving signal:noise ratio and visual performance unchanged (Meese *et al*, 2006). Interestingly then, binocular summation does not always lead to binocular advantage.

A similar relationship was seen in fMRI responses to grating contrast. At 2% contrast, BOLD responses to binocular input were notably larger than to monocular input, but at 10% contrast there was no difference in response, and this lack of additivity was attributed to interocular suppression or binocular contrast normalization (Moradi & Heeger, 2009).

A functionally important consequence of this balance between binocular summation and interocular suppression is *ocularity invariance*. Despite the marked difference in contrast thresholds, the perceived contrast of supra-threshold gratings is almost the same for one eye and for two eyes (Baker, Meese, & Georgeson, 2007; Ding, Klein, & Levi, 2013; Legge & Rubin, 1981). This form of perceptual constancy is likely to be important where the view of an object is partly obscured by a nearer one, such that part of the object's surface is seen by both eyes while the occluded part is seen by one eye (a 'half-occlusion'). Without ocularity invariance this switch in viewing conditions across the surface could be falsely taken as a change in contrast – a texture boundary - on the object itself.

Despite ocularity invariance, and the associated lack of binocular advantage in contrast discrimination, we found direct evidence that binocular summation occurs at all levels of contrast. The novel tactic here was to keep suppression almost constant by using a *binocular* pedestal grating of contrast C, and then to compare the detectability of monocular *versus* binocular contrast increments ΔC. A binocular advantage was revealed at all contrast levels C, because it was not offset by a corresponding increase in suppression (Meese *et al*, 2006).

Beginning with the pioneering work of Legge (1984), studies of this kind have aimed to make systematic and fairly precise measurements of contrast-difference thresholds over a wide range of binocular conditions, and from these increasingly rich datasets to construct and evaluate models for the functional architecture of signal-processing in binocular vision (Baker, Meese, & Summers, 2007; Baker, Meese, & Hess, 2008; Ding & Sperling, 2006; Hou, Huang, Liang, Zhou, & Lu, 2013; Huang, Zhou, Zhou, & Lu, 2010; Maehara & Goryo, 2005; Meese et al., 2006; Ding & Levi, 2016). Such models must specify the nature of the pathways from each eye, what the relevant signals are and how they interact, how the signals are combined, what and where the nonlinearities are, where the performance-limiting noise occurs, and how trial-by-trial perceptual decisions are made on

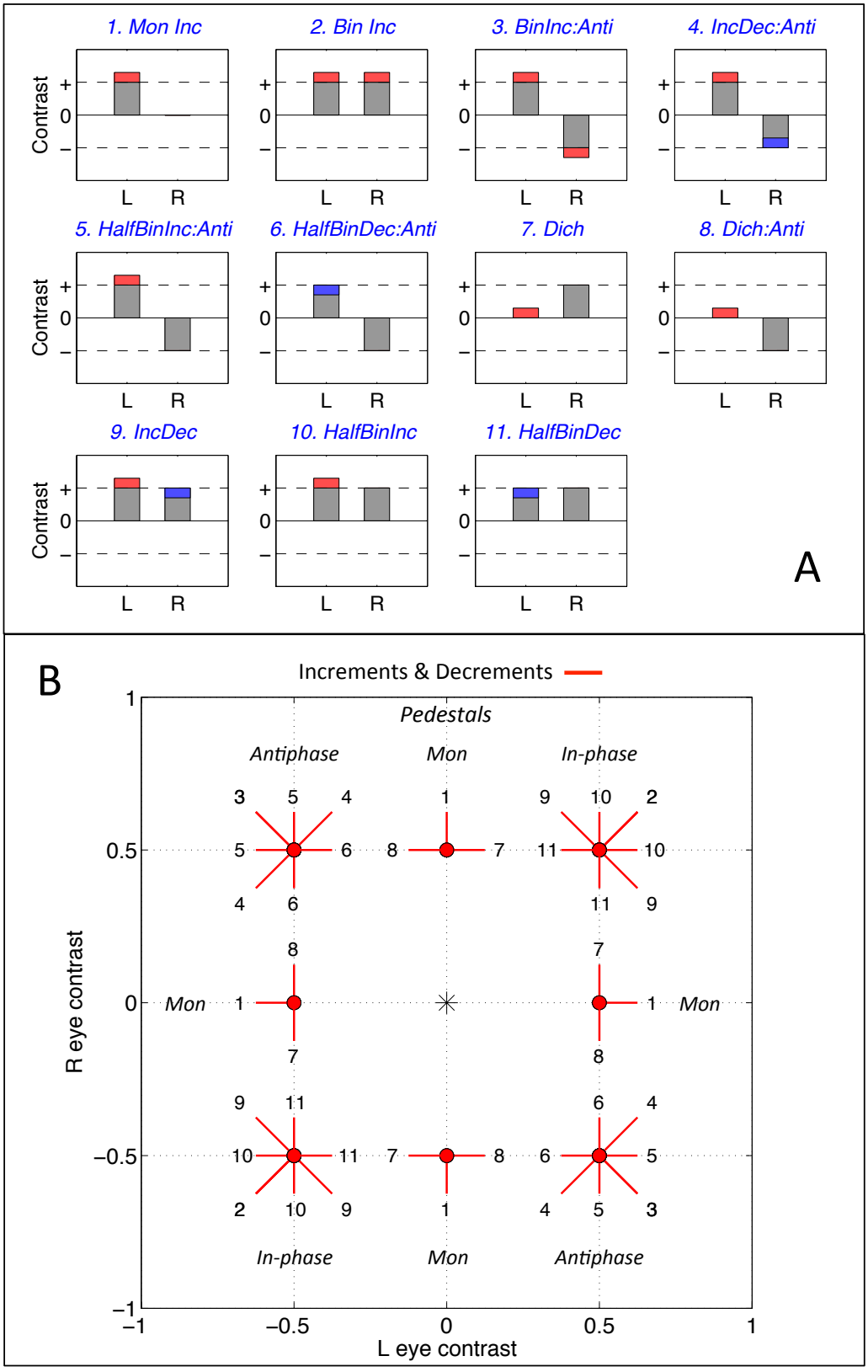

Figure 2. A: Graphical representation of the 11 different contrast discrimination tasks (Table 1). Pedestal contrast C (grey bars, dashed lines) may be increased (red) or decreased (blue) by some amount ΔC in the test interval. The observer's 2AFC task was to identify the test interval. B: Binocular contrast space. The 11 tasks can be seen as probing the visual system's response to changes in binocular contrast in the directions indicated by red lines, labelled with the corresponding condition numbers shown in panel A and Table 1. Red circles mark the pairs of pedestal contrasts (c_L, c_R); these could be in-phase, antiphase or monocular. The values of (c_L, c_R) were counter-balanced across left and right eyes, and across the sign of contrast (+ or -), and this led the 11 basic conditions to be reflected about the positive and negative diagonals, yielding a total of 40 distinct test vectors (red lines) for each pedestal contrast C. In our data analysis we assumed symmetry across the eyes, and across sign of contrast, and this reduced the number of different tasks back to 11.

the basis of one or more available outputs. Successful models for these contrast discriminations are likely to offer further insight into other binocular processes, such as binocular fusion, rivalry and stereoscopic vision.

In the present paper we extend the discrimination experiments of Meese, Georgeson & Baker (2006) with a set of critical new conditions that enable us to refine and expand our account of the functional architecture of human binocular contrast coding. The new experiments include conditions where (i) the target is a decrement of contrast rather than an increment, (ii) the target is an increment in one eye but a decrement in the other eye, and (iii) for each type of target, the pedestal gratings are out-of-phase ('antiphase') in the two eyes, rather than in-phase. Combining 6 new and 7 previous datasets gives us a total of 13 different discrimination functions (also known as *TvC* [threshold versus contrast] functions, or 'dipper functions') that need to be accounted for. The 13 functions comprise 11 distinct tasks, plus two replicates. This great variety of related discrimination tasks puts strong constraints on possible models of binocular signal processing. Put simply, we found that many models can fit data from some or even most of the eleven tasks; we found only one that accurately accounted for all eleven tasks at all contrast levels.

1.2 The discrimination tasks

The 11 tasks are defined schematically in Fig 2A. Grey bars represent the pedestal contrasts presented to one or both eyes; increments of contrast magnitude are shown in red, decrements of contrast magnitude in blue. Giving a short, unambiguous name to each task is not easy, but we have attempted to do so (see panel headings in Fig. 2A). The names can be cumbersome, so we rely a good deal on the numbering of tasks 1-11 throughout the paper, and invite the reader to decode the numbers via Fig. 2A.

It is also not easy to see much order or structure in the 11 conditions of Fig. 2A. The structure emerges clearly, however, when we consider the experiment in a two-dimensional *binocular contrast space*, whose axes are (c_L,c_R) - the contrasts shown to the left and right eyes (Fig. 2B). Monocular pedestals lie on the cardinal axes, binocular in-phase pedestals lie on the positive diagonal, and binocular antiphase pedestals lie on the negative diagonal (red symbols in Fig. 2B). Any change in (c_L,c_R) can be seen as a displacement from the pedestal point in some direction through this space. Red lines in Fig. 2B are *test vectors*, defining the direction of binocular contrast change for a given condition (1-11). For example, condition 2 (*BinInc*) has a binocular in-phase pedestal (top right in Fig 2B), and a binocular contrast increment that is an oblique displacement up and to the right. Condition 9 (*IncDec*) has the same binocular pedestal, but an increment in the left eye coupled with a decrement in the right eye, and this gives a test vector that points down and to the right. Counterbalancing across left and right eyes, and across absolute sign of contrast, reflects the 11 tasks about the positive and negative diagonals, yielding a total of 40 test vectors. This gives a pleasing symmetry to the experimental design, but more importantly it means that the set of pedestal positions and test directions gives a fairly comprehensive sampling of the discriminations that are possible in this space. This in turn puts strong constraints on the nature of binocular mechanism responses, and these can be expressed as *response surfaces* over the same space. The binocular contrast space (Fig. 2B) is a domain in which we can express the stimuli, the experimental design, the experimental results, and possible explanatory models.

1.3 Encoding contrast: monocular and binocular channels?

It is self-evident that the optic nerve fibre tracts leading from each eye are monocular pathways. Although the left- and right-eye layers of the LGN have the potential to interact with each other, and could be the earliest site for binocular rivalry (Haynes, Deichmann, & Rees, 2005), the earliest site for binocular summation appears to be the primary visual cortex. In most, perhaps all, quantitative models of binocular summation (Ding et al., 2013; Ding & Sperling, 2006; Legge, 1984; Maehara & Goryo, 2005; Meese et al., 2006; Meese & Hess, 2004), it has been tacitly assumed that only the binocularly-summed outputs are available to later stages of perception and decision, while the monocular pathways are not. They serve only as the input to binocular combination (Fig. 1a). The possibility of monocular outputs has been considered (eg. Legge, 1984), but to our knowledge there has been no critical discrimination experiment that would test for the availability of monocular outputs in parallel with the binocular ones (Fig. 1b).

Evidence from visual aftereffects suggests that it would be worthwhile to devise a rigorous test for monocular outputs. After adaptation through one eye, the tilt and motion aftereffects, and the contrast threshold elevation effect, can be observed when testing the same eye, and to a lesser degree when testing the other eye. The usual interpretation of such partial interocular transfer (Blake, Overton, & Lema-Stern, 1981; Moulden, 1980), and the finding of separate monocular and binocular motion aftereffects (Anstis & Duncan, 1983), is that distinct monocular and binocular neurons have been adapted. The monocular outputs might be separately available to perception, as in Fig. 1(b), but a single-output scheme (Fig. 1a) might also explain these aftereffects, provided the monocular input units were adaptable. In short, the involvement of monocular neurons in early visual coding seems very likely, but their functional

organization remains unclear. We aim to clarify these and other questions through discrimination experiments - a more incisive tool than adaptation.

1.4 The increment-decrement task

We define a monocular mechanism as one that is driven by contrast in one eye, but is unaffected by the other eye. A key test for the existence of separate monocular mechanisms is fairly straightforward. Suppose we perform a 2AFC discrimination task in which the non-target interval shows pedestal gratings of (say) 10% contrast to both eyes, while the target interval shows gratings of 12% to one eye but 8% to the other eye. The observer has to identify the target interval. If only the binocular summing mechanism exists (Fig. 1a), this task should be difficult or impossible because there should be little or no change in the binocular output between intervals. The extra response to contrast increment in one eye should be cancelled by the decrement in the other. Whether the cancellation was complete (making the task impossible) or partial would depend on the degree of nonlinearity in the contrast response before summation, but we can reasonably expect the contrast response function to be approximately linear over a narrow input range (eg 10±2 %), and so this increment-decrement task (henceforth *IncDec,* condition 9) should show poor performance from a binocular-summing mechanism. A monocular mechanism, on the other hand, should suffer no such difficulty because its response to the increment should, by definition, be unaffected by the decrement in the other eye (Fig. 1b). In short, performance on the *IncDec* task may reveal the presence, or absence, of specifically monocular mechanisms accessible to perception.

2 Methods

2.1 Visual display – conditions 1-6

Achromatic, horizontal, sinewave gratings of 1 c/deg were shown on a Clinton fast phosphor, high brightness CRT monitor at 120 Hz frame rate. The images were generated on a PC and displayed via a VSG interface card (CRS Ltd) which was synchronized with a pair of FE-1 ferro-electric shutter goggles (CRS Ltd) that enabled images to be shown separately to the left and right eyes on alternate frames. It is reasonable to consider such fast alternation of raster-scan images between the eyes to be effectively equivalent to a simultaneous steady presentation to each eye. At these frame rates (60 Hz per eye) no screen flicker is seen. Careful photometric measurements through the frame-interleaving goggles showed that the degree of 'crosstalk' between the eyes (the extent to which the left eye's image was visible to the right eye, or vice-versa) was negligible. The mean luminance of the display was 153 cd/m^2, but through the goggles this was attenuated by a factor of eight to 19 cd/m^2. A fixation point (dark dot, 2x2 pixels) was present throughout. Viewing distance was 57 cm, at which distance there were 28 pixels per deg of visual angle. The display luminance was gamma-corrected (linearized in relation to pixel greyscale values) using the CRS OptiCal photometer.

The gratings were defined by sinusoidal modulations of the mean luminance, restricted to a central window 5 deg in diameter (see insets to Fig. 3). The circular aperture W of the grating was smoothed by a 1 deg half-period of a raised sine-wave. This meant that a grating patch of contrast C was reduced to contrast $C/2$ at a radius of 2 deg, and to zero at 2.5 deg. This reduced truncation artefacts – sharp edges that might be a spurious cue to detection. Thus the horizontal grating $L(x,y)$ was defined by its modulation of the mean luminance L_0:

$$L(x,y) = L_0\{1 + W(x,y).(C \pm \Delta C)\sin(2\pi f y - \phi)\}$$

where f is spatial frequency, ϕ is phase in radians, C is the pedestal contrast magnitude, and ΔC is the change in contrast that defines the target interval on each trial. Phase relative to the screen centre (y=0) was the same in both intervals of a given trial, but varied randomly [ϕ = 0, π/2, π, or 3π/2] across trials. We must make a clear distinction between *phase* and *polarity*. Whatever the chosen spatial phase value, the pedestal grating might have the *same* polarity in each eye (C,C) or the *opposite* polarity ($C,-C$). For consistency with previous work, we also refer to the same-polarity conditions as *in-phase*, and the opposite-polarity conditions as *antiphase*. Finally, for any phase and polarity, the target grating might be defined by a contrast *increment* ($C+\Delta C$) or a *decrement* ($C-\Delta C$). An increment in one eye might be accompanied by an increment, a decrement or no change in the other eye. Table 1 lists the full range of 11 different tested conditions defined in this way; conditions 12 and 13 replicate conditions 1 and 2. For decremental targets we ensured that $\Delta C <= C$, so that variations in ΔC never entailed a reversal of grating polarity.

2.2 Procedure – conditions 1-6

Detectability of the contrast difference ΔC was assessed with a 2-interval forced-choice staircase method. The task was to identify which of two 200ms presentations, defined by audible tones and separated by a 500 ms blank (mean luminance) period, contained the contrast difference ΔC. To enable learning and to encourage best performance, auditory feedback about correctness (a high or low tone) was given after each trial. Trials for a given condition (numbers 1-6; table 1) were tested in separate sessions, and different pedestal contrasts C were tested in different blocks of trials within a

session. Sessions and blocks were randomly ordered. The staircase rule reduced contrast by 1 step after 3 correct trials, and increased it by 1 step after each incorrect trial. The step size within each block was initially large (8 dB) but reduced to 4 dB after the first reversal of staircase direction, and then to its final value of 2 dB after the second reversal. Each staircase ran for 50 trials.

Observers were two of the authors (SAW, DHB), who had much previous experience of contrast detection and discrimination experiments, and a third less practised observer (ASB). Informed consent was obtained and the work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). All observers were given about ten minutes practice in condition 2 before starting the experiment. We did not attempt to define the subjective impressions that might be created by the various dichoptic test conditions. Instead, observers were instructed to respond so as to maximise the number of correct feedback tones and were not informed about which condition was being tested in a given session. For every left eye/right eye condition defined in table 1, we also ran a corresponding right eye/left eye condition and the data were pooled to average out any ocular asymmetries, which generally appeared to be small, perhaps because the spatial frequency (1 c/deg) was not high.

2.3 Analysis – conditions 1-6

For each observer, raw data were pooled across 5 repeated sessions, and across corresponding left eye/right eye conditions. Psychometric functions (cumulative Gaussians, defined by proportion correct as a function of log(ΔC)) were fitted by probit analysis, and each fitted function was summarized by its threshold value – the contrast required to achieve 75% correct. Thus each threshold was derived from a total of 500 trials.

2.4 Analysis – conditions 7-13

Thresholds for conditions 7-13 were drawn from our previous studies, as listed in table 1. Procedure was very similar to that described above, though some details differed. The main factors – 2AFC staircase procedure, grating orientation and spatial frequency, grating patch size, pedestal contrast levels – were common to all conditions. Conditions 9-13 used a mirror stereoscope instead of stereo goggles, and 100ms duration instead of 200ms. The published thresholds were defined at 81.6% correct (d'=1.3) on a fitted Weibull function, but thresholds were re-computed here for 75% correct to match all the other conditions, and are shown as the geometric means of the two observers (MAG, TSM). Conditions 7 and 8 report the geometric mean thresholds for 2 observers (DHB, LP) again at 75% correct.

Table 1: Pedestal (C) & Test contrast (ΔC) relations in the 13 conditions

Cond No.	*Fig.4 panel*	*Source*	*Condition name*	Test interval		Non-test		*Type of pedestal*	*No. of Ss*	*Dur, msec*
				L eye	*R eye*	*L eye*	*R eye*			
1	A	New	MonInc	C+ΔC	0	C	0	Monocular	3	200
2	A	New	BinInc	C+ΔC	C+ΔC	C	C	Binocular	3	200
3	B	New	BinInc Anti	C+ΔC	-(C+ΔC)	C	-C	Antiphase	3	200
4	B	New	IncDec Anti	C+ΔC	-(C-ΔC)	C	-C		3	200
5	C	New	HalfBinInc Anti	C+ΔC	-C	C	-C		3	200
6	C	New	HalfBinDec Anti	C-ΔC	-C	C	-C		3	200
7	D	B&M	Dich	ΔC	C	0	C	Dichoptic	2	200
8	D	B&M	Dich Anti	ΔC	-C	0	-C		2	200
9	E	M,G&B*	IncDec	C+ΔC	C-ΔC	C	C	Binocular	2	100
10	E	M,G&B	HalfBinInc	C+ΔC	C	C	C		2	100
11	E	M,G&B*	HalfBinDec	C-ΔC	C	C	C		2	100
12	F	M,G&B	MonInc	C+ΔC	0	C	0	Monocular	2	100
13	F	M,G&B	BinInc	C+ΔC	C+ΔC	C	C	Binocular	2	100

Notes: Reference to left and right eyes is nominal; all conditions were counter-balanced across left and right eye presentation. B&M = Baker & Meese (2007); M,G&B = Meese, Georgeson & Baker (2006); *unpublished data from the study of M,G&B.

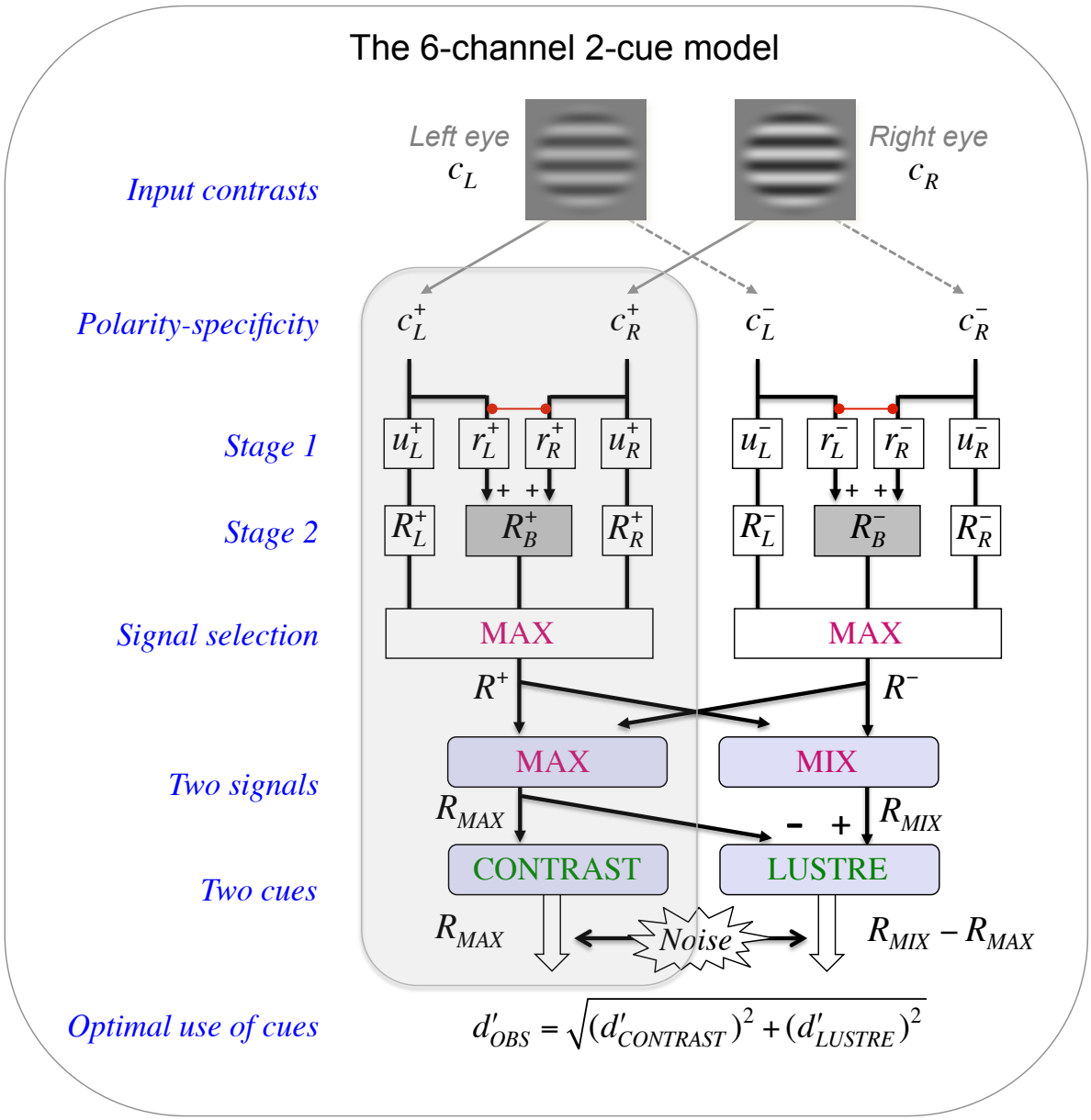

Figure 3. Levels of processing in the model architecture, from the pair of input contrasts (c_L,c_R) to observed discrimination performance (d'_{OBS}). Subscripts *L,R,B* denote Left eye, Right eye, and Binocular respectively. The binocular response R_B is the same as in our earlier model (Meese *et al*, 2006) and incorporates both interocular suppression (red links) at Stage 1 and a nonlinear response function at Stage 2 (not illustrated, but see Appendix 1 for model equations). Superscripts '+' and '-' denote separate responses to stimuli of opposite contrast polarity (e.g. gratings of opposite phase). For in-phase pedestals (e.g. both '+') the model is relatively simple; only the positive-polarity pathway (shaded) leading to the *contrast* cue needs to be considered. With inputs of opposite polarity to the two eyes, *lustre* is assumed to be a second possible cue. The two cues, *contrast* and *lustre,* are perturbed by late noise. The observer is assumed to make use of both cues to perform the discrimination task.

3 Model development

In a well-established tradition of modelling (Foley, 1994; Legge, 1984; Legge & Foley, 1980; Legge, 1979), the present model greatly extends our earlier 2-stage binocular channel model (Meese et al., 2006) particularly by (i) introducing separate mechanisms for opposite contrast polarities, not previously considered in 2006, but introduced by Baker & Meese (2007), (ii) introducing monocular mechanisms in parallel with the binocular-summing channel, and (iii) introducing signal selection and decision rules based on the Minkowski sum, and the MAX operator, to handle the multiplicity of outputs. The rationale for these key modifications will unfold as we analyze and fit models to the results, but first we describe informally the architecture (Fig. 3) that emerged as most consistent with the full pattern of our findings. Appendix 1 gives the full set of equations that define the model's responses and behaviour.

3.1 Signal processing stages 1 and 2

We assume polarity-specific signals right from the start. Responses in each eye proportional to retinal contrast are carried by separate channels for positive and negative sign of contrast (Fig. 3). The likely neural basis for such a separation is the division of retinal ganglion cells into two classes, ON-centre and OFF-centre.

The binocular channel, with monocular inputs r_L, r_R and combined output R_B, is essentially identical to the 2-stage model of Meese *et al* (2006), but with polarity-specificity now made explicit. As in Meese *et al* (2006), the two monocular responses r_L, r_R are subject to ipsiocular and interocular suppression (contrast gain control), and the output R_B is subject to a form of smoothed thresholding (determined by the value of the saturation constant, z), and a power-law nonlinearity at higher response levels (determined by the difference between exponents in the numerator and denominator, p-q).

A major new addition is the introduction of truly monocular channels in parallel with the binocular ones. These have inputs u_L (or u_R), with output R_L (or R_R), and are the same as the binocular channel in all respects *except* that any influence from the other eye at stage 1 and stage 2 is deleted. This proves to be a useful and parsimonious assumption that introduces no new parameters. It implies that for a monocular input image (e.g. in the left eye), $R_L = R_B$.

3.2 Signal selection

The two triplets of *L,B,R* channels thus create 6 signals (3 for each polarity) that need to be dealt with. There are many possibilities. In the main experiment, and in pilot experiments, we found that with a binocular in-phase pedestal, detecting a decrement of contrast in one eye (condition 11) was very much harder than detecting a similar increment in one eye (condition 10). For example, if the [L,R] pair of pedestal contrasts (in %) was [10, 10], then discriminating that from the monocular decrement [8, 10] was much more difficult than for the corresponding increment [12, 10]. This reliable finding prompted the idea that the visual system might simply use the largest of the L,B,R signals - a *MAX* operator - rather than combining them in any more substantial way (Fig. 1c). As a useful intuition, note that there's no difference between *max*(10,10) and *max*(8,10) so a *MAX* operator applied to the pair of monocular contrasts could never detect the monocular decrement, but would detect the monocular increment (the *max* increases from 10 to 12). But we must also consider the B channel. Suppose the B channel *averaged* the monocular contrasts, and then one signal was selected as *max*(L,B,R). The selected pedestal response then becomes *max*(10,10,10), to be compared with the incremental case *max*(12,11,10), versus the decremental case *max*(8,9,10). The outcome is unchanged from the 2-channel example. Thus polarity-specific signal selection, $R = max(R_L, R_B, R_R)$ was a plausible candidate to be added to the output of stage 2. This reduced the six responses to two: R^+ and R^-, representing the positive and negative contrasts *irrespective of ocularity* (Fig. 3). For monocular or in-phase inputs, one of R^+ or R^- would always be zero; but for antiphase inputs, both would be active. This presents a potential conflict that we propose is resolved by creating *two* perceptual cues that are the subjective basis for discrimination: *contrast* and *lustre*.

3.3 Contrast and Lustre

The simultaneous presence of opposite polarities has been studied a good deal in the context of stereo vision, binocular brightness and contrast perception, binocular summation and binocular fusion. There is little or no binocular advantage for antiphase signals at contrast threshold (Green & Blake, 1981), no evidence for binocular fusion (single vision) for antiphase signals (Georgeson & Wallis, 2014; Schor, Wood, & Ogawa, 1984), and no sense of stereo depth when one image of a random-dot stereo pair is reversed in contrast (Cumming, Shapiro, & Parker, 1998; Julesz, 1971). This large literature on the perception of opposite polarities (also see Howard & Rogers, 1995) supports our assumption above that binocular summation (in the B channels) is polarity-specific, but also has two other implications.

First, in earlier work we found that binocular antiphase gratings could appear to have about the same perceived (matched) contrast as in-phase gratings (Baker, Wallis, Georgeson, & Meese, 2012). But at lower standard contrast levels antiphase contrasts appeared lower than in-phase, and more similar to monocular gratings. Zhou, Georgeson, & Hess (2014) showed that these contrast-matching data, and spatial phase-matching data, were well explained by a model that included a noisy *max* over monocular and binocular response amplitudes, regardless of their spatial phase. We adopt the same idea here, using $max(R^+, R^-)$ as the code for contrast (R_{MAX} in Fig. 3). It implies that dichoptic *contrast* perception is determined by whichever polarity has the larger response at a given time.

Second, we should consider a possible contribution from the opposite polarity. *Lustre* is a kind of shiny, metallic appearance that often arises when opposite polarities are shown to the two eyes, or rapidly flickered over time in one eye (Anstis, 2000; von Helmholtz, 1925). For references to 19th and 20th century research, see Bixby (1928), Mausfeld, Wendt, & Golz (2014). Wolfe & Franzel (1988) found that visual search for a lustrous target amongst non-lustrous distractors was rapid and

independent of distractor set size, implying parallel search, and suggesting that lustre, like contrast, might be a basic feature in early vision. Interestingly, search for rivalrous targets did not have these characteristics, suggesting that rivalry is not a basic feature. Lustre may be "vision's response to two conflicting signals from one region of the visual field" (Anstis, 2000) and it could be a second cue for discrimination in those of our tasks that involve antiphase gratings.

How should we model the response to lustre? We begin with the working hypothesis that lustre is, in some manner, a perception of light and dark at the same time. This is consistent with many earlier observations (von Helmholtz, 1925), and subjective descriptions (e.g. Bixby, 1928) and previous experiments (Anstis, 2000). It suggests that in the model we should create a signal that pools over (R^+, R^-). Suppose, for example, that this signal (R_{MIX}) was the quadratic sum of (R^+, R^-). R_{MIX} would pool responses to opposite polarities, but it would also respond to non-lustrous inputs (R^+ or R^- alone). To create a more specific response to lustre (R_{LUSTRE}), we need to remove the non-lustrous component, and this idea suggests a general formulation (Fig. 3): $R_{LUSTRE} = R_{MIX} - R_{MAX}$. The nature of R_{MIX} remains to be determined, but with the requirement that for non-lustrous inputs $R_{MIX} = R_{MAX}$.

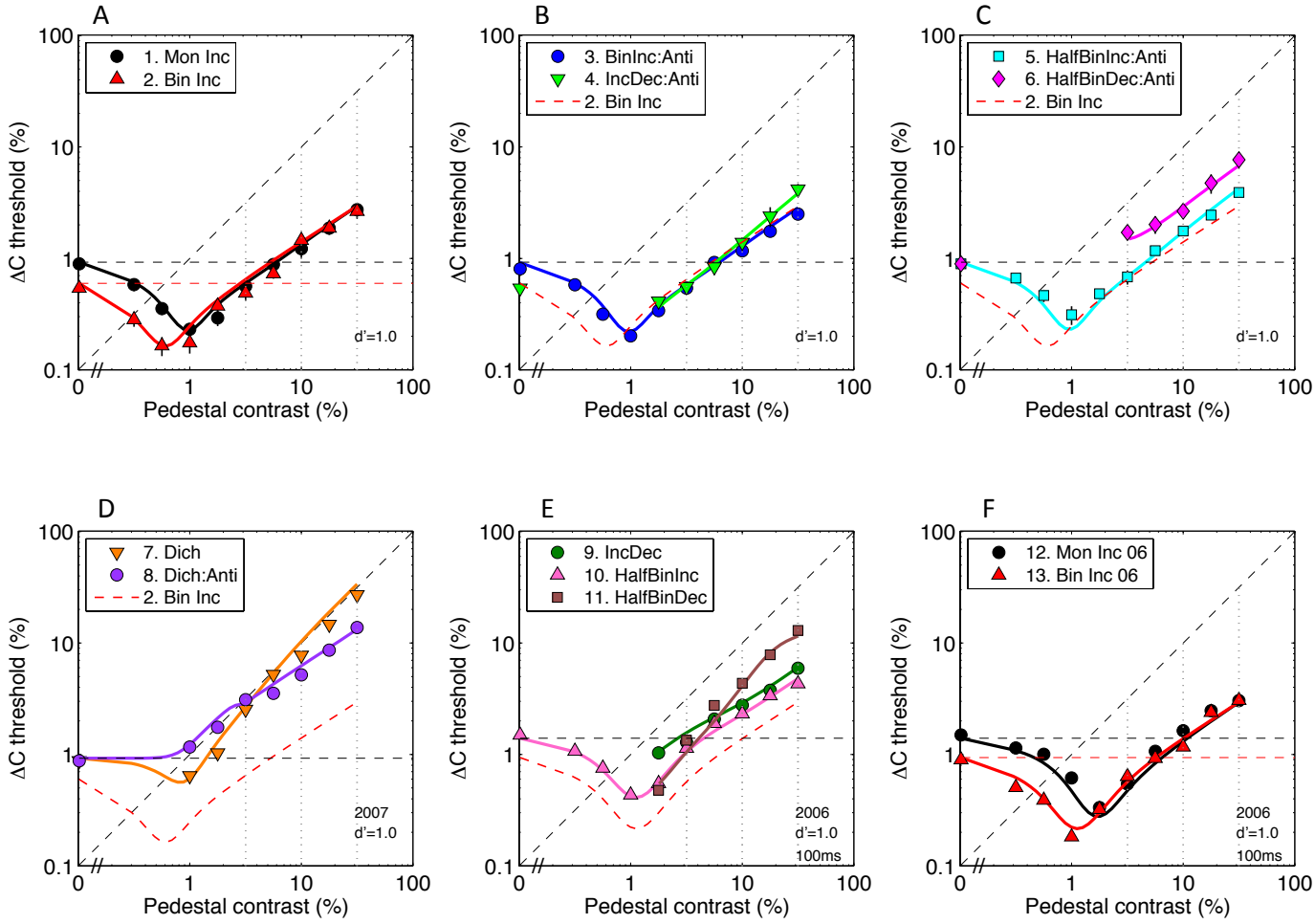

Figure 4. Experimental discrimination thresholds (symbols) and model fitting (curves). A-C: new data (conditions 1-6); geometric mean thresholds across 3 observers. D-F: data re-plotted from our previous studies (conditions 7-13, see Table 1); geometric means across 2 observers. Model fit was excellent: RMS error = 1.16 dB; R^2 = 0.984. Black and red horizontal dashed lines are the model's monocular and binocular contrast thresholds respectively. Oblique dashed line is the locus of points where ΔC = C; it is the upper limit of testable ΔC values for contrast decrements (conditions 4,6,9,11). Binocular contrast increment detection (condition 2 or 13, red curve) is a useful baseline against which to judge other conditions. This baseline is copied into the other panels as a red dashed curve (curve 2 into panels B,C,D; curve 13 into panel E).

4 Results

4.1 'Dipper' functions

Results from the 11 different monocular, binocular and dichoptic contrast discrimination tasks (Fig. 2A) are summarized in Fig. 4 as log-log plots of just-discriminable contrast change ΔC *versus* pedestal contrast C. Thirteen separate 'dipper functions' are shown, because conditions 1,2 replicated conditions 12,13 (see Table 1). Some key features of the results in Fig. 4 are:

• For in-phase gratings, binocular advantage occurred only at low contrasts. Binocular contrast discrimination thresholds (condition 2) were lower than monocular (condition 1) only for low or zero pedestal contrasts. This is not new, but it reinforces the soundness of similar findings by Legge (1984), Maehara & Goryo (2005), and Meese *et al* (2006) whose data are re-plotted here as conditions 12, 13. There was no binocular advantage when the pedestal was visible, above about 1% contrast. And yet, when the pedestal was binocular, thresholds for increments in one eye (condition 10, termed 'half-binocular' increments) were about a factor of 2 (mean 5.2dB) higher than for binocular increments on the same binocular pedestal (condition 13), implying binocular summation across the whole range of contrasts. These results together imply that binocular summation can confer a binocular advantage, but does not always do so (cf. Meese et al, 2006).

• For antiphase grating detection without a pedestal, there was only a very small binocular advantage over monocular detection (mean 0.89 dB), consistent with previous studies. Binocular summation is evidently phase- or polarity-specific (Cogan, 1987; Cohn & Lasley, 1976; Green & Blake, 1981), and that rules out binocular energy summation as a candidate mechanism, since that would show binocular advantage for antiphase as well as in-phase (Westendorf & Fox, 1973).

• For antiphase pedestal gratings, thresholds for discriminating an increase of contrast in both eyes (condition 3) were equal to those for increments on a monocular pedestal (condition 1), even with low contrast pedestals. The mean threshold difference across all pedestal levels was tiny and insignificant (0.35 dB). It is perhaps surprising that, in the face of possible rivalry, antiphase discrimination thresholds were not *higher* than corresponding monocular thresholds. It could be that antiphase signals simply fail to sum, but do not cancel each other out. This could be achieved by half-wave rectification before summation. Alternatively, antiphase signals might cancel each other in a binocular summing mechanism, while performance is carried by monocular channels in parallel with the binocular ones (Fig. 1b). These questions cannot be answered from the data alone, but can be addressed by modelling.

• Despite the lack of antiphase binocular advantage just described (condition 3 *vs* 1), we found that thresholds for antiphase binocular increments (condition 3) were moderately but systematically better than the corresponding half-binocular increments (condition 5), by an average of 2.7 dB, perhaps implying some weak form of antiphase summation.

• If only the binocular summing mechanism B existed (Fig. 1a) then we should make two predictions, both of which turn out to be contradicted by the data. (i) For in-phase pedestals, it seems likely that the combined increment-decrement (condition 9) should be especially hard to detect, because opposite changes would cancel in the binocular sum or binocular average (see *Introduction*). In the experiment, this was not so; thresholds for condition 9, averaged over the 4 pedestal contrasts higher than 5%, were only slightly (1.5dB) higher than for the 'half-binocular' increment (condition 10). (ii) Conversely, it is reasonable to expect increments and decrements to be about equally detectable by the B mechanism, but in fact 'half-binocular' decrements (condition 11) were much harder to see than 'half-binocular' increments (condition 10). Thresholds were a factor 2 (6.4dB) higher than for the corresponding increments (condition 10) (again averaged over the 4 highest pedestal contrasts). In short, the condition that should be easy for the B mechanism is difficult, and vice-versa. These two results imply that, even for in-phase gratings, the B mechanism is not alone. Parallel monocular mechanisms seem likely. But in highly nonlinear models even simple intuitions of this kind can be misleading or depend heavily on other unrecognized assumptions. To draw firmer conclusions, we need to go beyond intuition and be guided by more precisely formulated, testable models.

4.2 Model fitting

Discrimination thresholds in dB (means of 2 or 3 subjects; Table 1) from all 13 dipper functions (11 different tasks, N=111 data points) were

fitted in the same run of the model, with 9 free parameters. RMS error was 1.16 dB, $R^2 = 0.984$, an excellent overall fit. Best-fitting parameters are given in Table 2.

Curves in Fig. 4 are the threshold curves generated by the best-fitting model. The match between model and data is strikingly good across the whole dataset, with no local anomalies. This is important, because a low RMS error could occur when (say) 9 of the 11 tasks fit very well, but two fit poorly. It was not difficult to find and reject models of that kind, and not easy to find the one that fitted well everywhere. The functional architecture (Fig. 3) and the parameter values (Table 2) are both of great importance. And yet there is still an explanatory gap: we need to understand *how* the proposed mechanisms and processes lead to correct predictions about the observer's behaviour. To do this we interrogate the model in revealing ways, by representing mechanism responses in *binocular contrast space*.

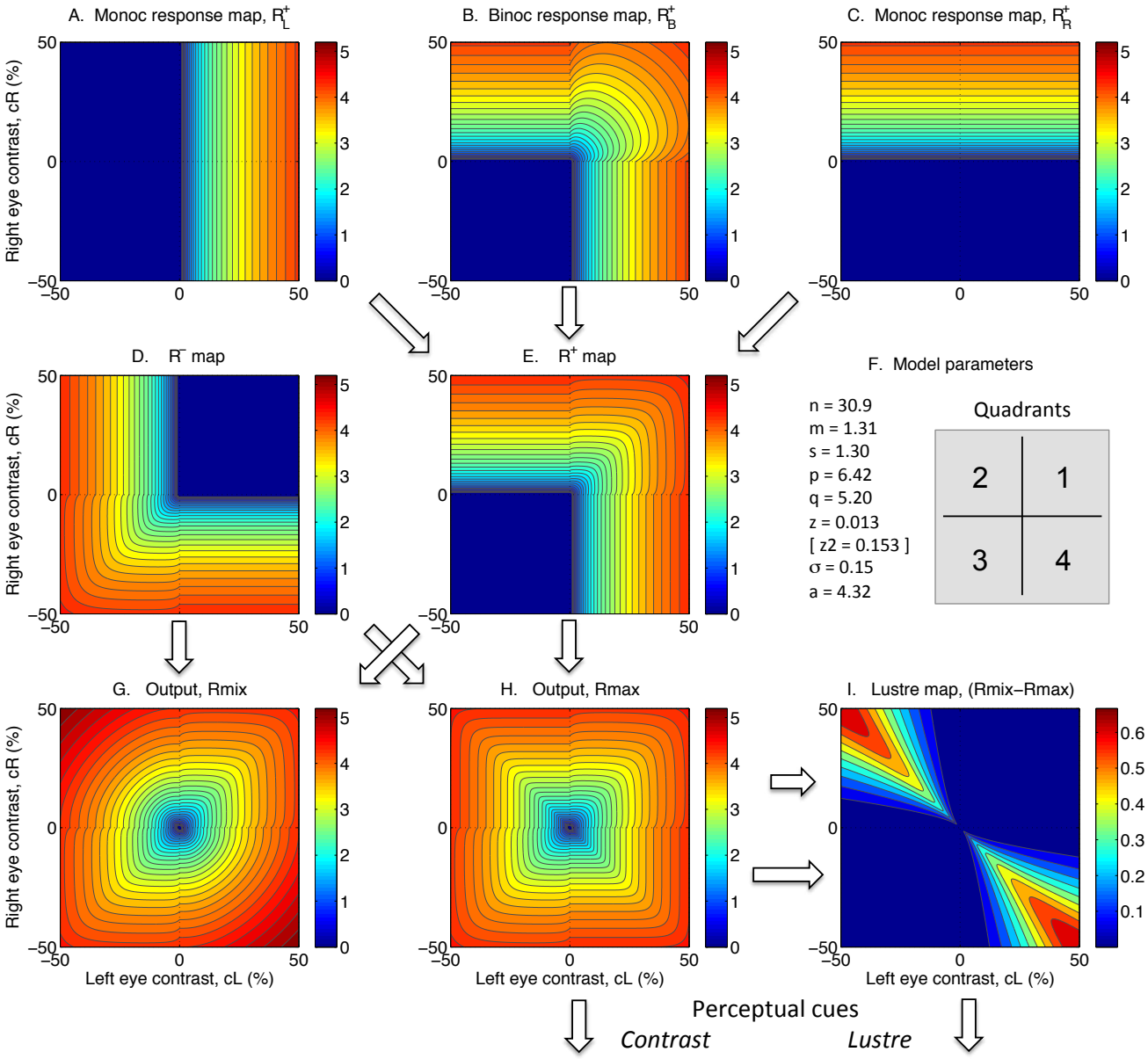

Figure 5. Each row represents responses at different stages of the model, mapped over the binocular contrast space. A,B,C: the three response maps from stage 2 (Fig. 3), representing positive contrast polarity in the channels for left-eye, binocular and right-eye respectively. E: These three maps are combined via the MAX-like operator, to create the *R*+ map. D: The corresponding L,B,R maps for negative polarity (not shown) are combined to form the *R*- map. H: The responses *R*+, *R*- are similarly MAX-ed to create the output cue R_{MAX} which we associate with the perception of contrast. G: The maps *R*+, *R*- are pooled again, in a way that is less MAX-like in the 2nd and 4th quadrants, to form R_{MIX}. I: The second perceptual cue, R_{LUSTRE}, is formed as the difference between panels G and H; $R_{LUSTRE} = (R_{MIX} - R_{MAX})$.

Table 2: Parameters used in the fitted model

n	30.914
m	1.31356
s	1.29675
p	6.41616
q	5.19607
z	0.01297
σ	0.14873
a	4.3227
z_2	0.15281

Note: the model in principle has 8 free parameters. The ninth parameter ($z2$) substitutes for z in conditions 9-13 only, for pragmatic reasons explained in Appendix 1 (Fitting the model).

4.3 Model behaviour: mapping the binocular contrast-response surfaces

Mechanism responses were computed from the model equations (Appendix 1) using the best-fitting parameters (Table 2), and visualized as 3D surfaces in binocular contrast space (Fig. 5). That space can be divided into four quadrants, defined in Fig. 5F. The monocular response R_L^+ (Fig. 5A) increases with positive left-eye contrasts (first and fourth quadrants), but is insensitive to any right-eye contrasts. Its right-eye counterpart (R_R^+) is equivalent, but rotated by 90^0 (Fig. 5C). The binocular channel (R_B^+) shares one quadrant with each of the monocular channels, and (by design) shows binocular interaction only in the first quadrant where both contrasts are positive (Fig. 5B). The R^+ response (Fig. 5E) can be envisaged as the envelope of these three surfaces. Note how its surface shape in the first quadrant differs from all three of the input surfaces. The R^- map (Fig. 5D) is a reflection of R^+ about the negative diagonal. R^+ and R^- are combined in two ways to form R_{MIX} and R_{MAX} (Figs. 5G, 5H). R_{MIX} and R_{MAX} are identical in the first and third quadrants, but differ markedly in the second and fourth (opposite-sign) quadrants, where R_{MIX} shows substantial, roughly quadratic, combination of left- and right-eye contrasts, while R_{MAX} is close to winner-take-all. This difference creates the lustre response (Fig. 5I), present only in the opposite-sign quadrants.

4.4 In-phase pedestals

The value of these maps should now become clear as we show how model predictions and observed discrimination thresholds can be understood and compared directly on the model response surface. Fig. 6A shows the 1st quadrant of the R_{MAX} surface in grey, and several iso-height contours (lines of constant response) are highlighted. White points represent the pedestals, and the outermost black curve represents the response level evoked by the highest pedestal contrast (31.6%). In signal detection theory, it follows from the definition of d' that to be just-distinguishable from the pedestal, *any* test condition must evoke a mean response that is one standard deviation (σ) higher or lower than the mean pedestal response. The locus of all such threshold points is therefore the pair of surface contours (red in Fig. 6A) whose height is σ above or below the pedestal contour (black). If and when R_{MAX} is the cue (decision variable) used by the observer, then observed thresholds should lie on these contours. More precisely, they should lie at the intersection of the test vectors (white) and the threshold contours (red). Red symbols in Fig. 6A represent observed thresholds on each test vector for conditions 9,10,11,13, and it is clear that they lie very close to the model's threshold contours. Similarly good agreement between model and data holds for the lower pedestal contrasts shown in Fig 6A (orange and green curves), and the even lower pedestals plotted in Fig. 6B. The surface contours change shape as contrast is reduced, but the data hug the model curves about equally well at all contrast levels.

This representation of data and model in binocular contrast space reveals a functional relation between different test conditions that is not evident from the 'dipper functions' alone. For example, returning to the highest pedestal contrast (red in Fig 6A), we can see that thresholds for condition 9 (*IncDec*) are higher than condition 13 (*BinInc*) because the test vector for condition 13 takes the shortest route to the threshold contour, while in condition 9 it passes rather obliquely across the surface, and so requires a greater contrast change to reach the same contour. We can also infer that condition 9 was detected as an increment, like conditions 10 and 13, but the decrement in the other eye shifted the direction of change, and made the task harder. Condition 11 (monocular decrement on a binocular pedestal) was even more difficult because the test vector ran almost parallel to the surface contours, rather than across them, and so much greater contrast change was needed to reach the threshold contour. We can also infer that condition 11 was seen as a decrement in contrast, not an increment. Threshold for the corresponding increment (condition 10) was almost 10dB lower at this pedestal contrast, because its test vector enjoyed a much more direct route to the threshold contour.

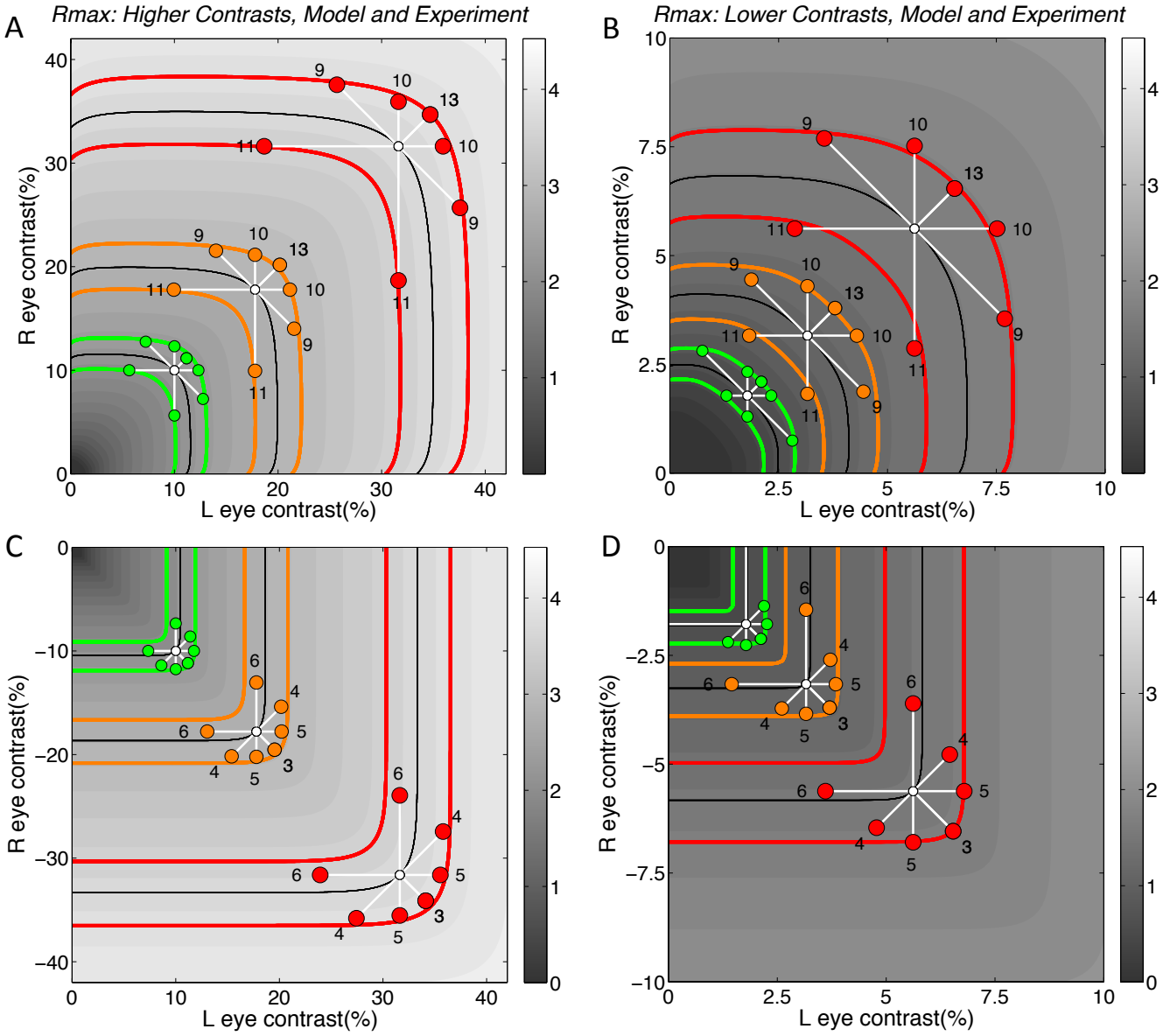

Figure 6. How the response surfaces (Fig. 5) can be used to understand and predict discrimination performance. A,B: Close relation between the 1st quadrant of the model output surface R_{MAX} (Fig. 5H; rendered in grey here) and the discrimination thresholds for in-phase pedestals (where $c_L = c_R$, conditions 9,10,11,13). A: White points represent the 3 highest pedestal contrasts (10.0, 17.8, 31.6%), each surrounded by a cluster of 7 data points (4 independent points, plus 3 mirrored across the positive diagonal) that represent the pairs of L,R contrasts that are just discriminable from the binocular pedestal in each of the 7 test directions (Fig. 2B). Thin black curves are the 3 iso-response contours of the R_{MAX} surface that pass through the 3 white pedestal points. Each pair of coloured curves (green, orange, red) represents the locus of all just-discriminable (d'=1) responses that lie one noise unit (σ) above, or below, the corresponding pedestal response level (black curve). If the model is correct, the observed discrimination thresholds (green, orange, red circles) should lie on or close to these curves. A very close fit is observed. B: As panel A, but zoomed-in to low contrasts, illustrating data for 3 lower pedestal contrasts (1.8, 3.2, 5.6%). C,D: As panels A,B, but for the 4th quadrant of the R_{MAX} surface, illustrating results for antiphase pedestals (where $c_R = -c_L$, conditions 3,4,5,6). Thresholds in conditions 3,4,5 fell close to the surface contours predicted by R_{MAX} but for condition 6 they did not. Condition 6 best reveals the perceptual contribution made by lustre rather than contrast - see Fig. 7.

At lower pedestal contrasts (Fig 6B) the surface contours change shape, exhibiting a wider range over which roughly linear summation of contrasts occurs (implied by approximately left oblique surface contours), and as a result the thresholds for conditions 10 and 11 become much more nearly equal. Interestingly, there are some test directions lying between conditions 9 and 11 that must be expected to have immeasurably high thresholds, at any pedestal contrast, because their test vectors would never intersect the threshold contour.

So far then, we have seen that the discrimination thresholds plotted in binocular contrast space give surprisingly direct information about the 3D shape of the response surface - including both the shape of surface contours and their vertical spacing in a fairly wide neighbourhood around each pedestal point. This conclusion should be sound when only *one* cue - hence a single response surface - is involved in the task. But we now turn from in-phase to antiphase pedestals, where *two* cues appear to contribute to performance.

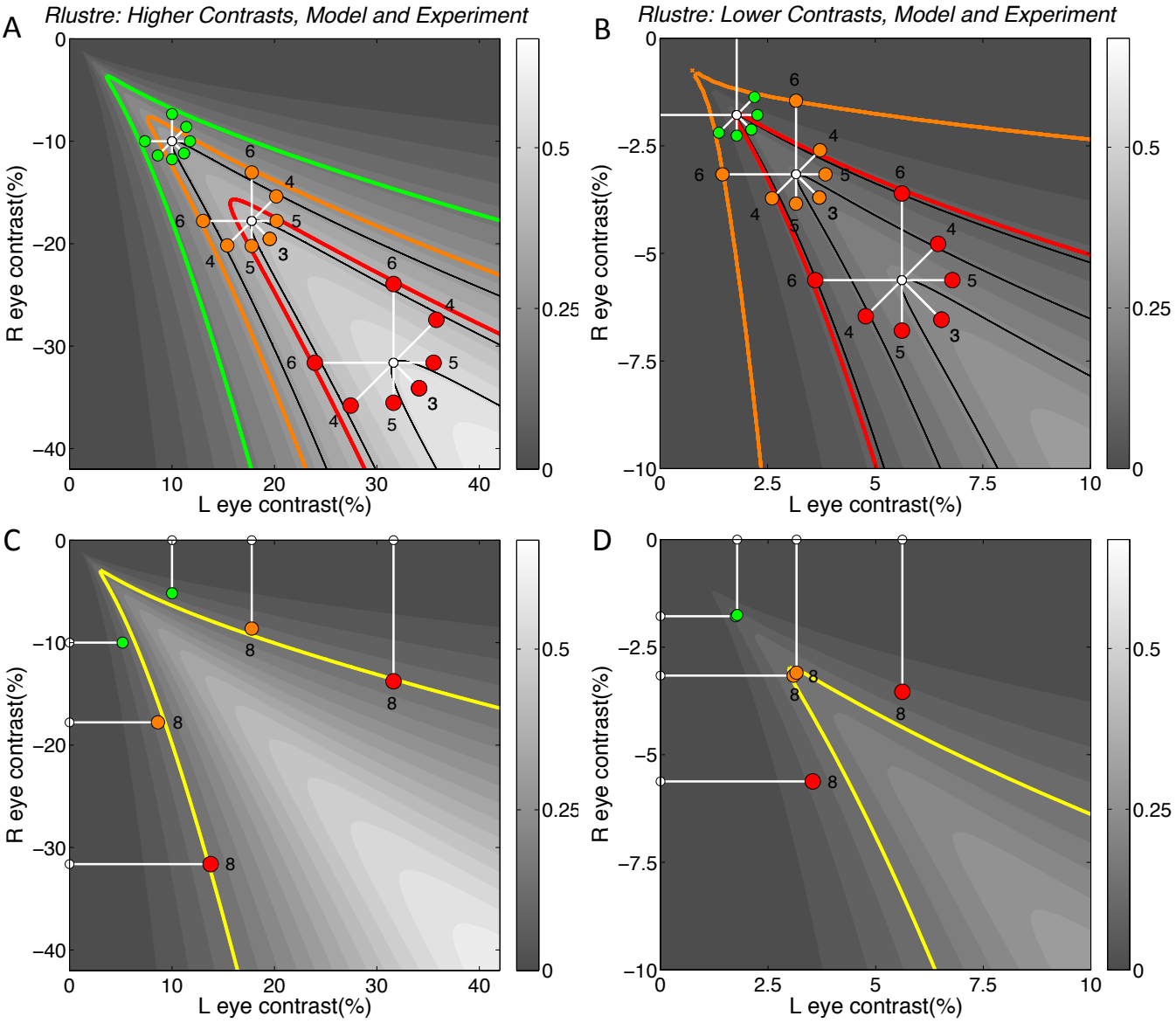

Figure 7. Antiphase conditions: how *lustre* contributes to performance in conditions 4, 6, 8, but not conditions 3 or 5. A,B: Plotting conventions and experimental data are as Figure 6C,D, but the model surface is R_{LUSTRE} rather than R_{MAX}. Note: Coloured contours here represent a predicted discriminable decrease of lustre. Decreases were important. Threshold contours for discriminable increases of lustre are not shown, because for antiphase pedestals any increase of lustre (condition 3) was too small to make any practical contribution to observed performance. Thresholds for condition 6 are close to the surface contours predicted by lustre at all 3 contrasts (green, orange, red) in A, and at 2 of the 3 lower contrasts (orange, red) in B. At the lowest pedestal contrast in B (green, 1.8%) the lustre cue was too weak to generate a threshold contour, and this was matched by absence of a reliable experimental threshold for condition 6 at this contrast level. Thresholds for condition 4 were close to the *lustre* predictions at higher contrast (A) but less so at lower contrast (B). In conditions 3 and 5, thresholds were markedly better (lower) than predicted by lustre. C,D: An *increase* of lustre explains performance in test condition 8 (pedestal in one eye, antiphase test grating in the other eye), except at very low contrast. Pedestal points (white) now lie on the c_L or c_R axes, where R_{LUSTRE} =0. Increasing test contrast ΔC increases lustre (along the white lines). Observed discrimination thresholds (coloured symbols) were mostly very close to the model's threshold contour for detecting lustre (yellow curve, defined by d'=1 and R_{LUSTRE} =σ; same contour for all the pedestal contrasts of condition 8). At the lowest pedestal contrast (1.8%, green, panel D) lustre was again too weak to be detected.

4.5 Antiphase pedestals and the lustre cue

Figs. 6C and 6D present the same form of analysis as Figs 6A,B, but for the antiphase pedestals. Threshold points for conditions 3,4,5 fell close to the R_{MAX} contours at all six pedestal levels. This consistency of shape (the 'rounded square' corner) strongly implies that the surface shape does not change much with contrast level. But thresholds for condition 6 (monocular decrement) consistently failed to fall on the predicted contours. This deviation is most obvious at the lower pedestal contrasts (Fig. 6D), but even at the higher contrasts (Fig. 6C) we should emphasize that the test vectors for condition 6 run *parallel* to the threshold contours, and so no adjustment of ΔC could take the threshold points any closer. The R_{MAX} surface thus predicts an immeasurably high threshold at all pedestal levels in condition 6, but the observed thresholds were only 5-6dB above their incremental counterpart (condition 5): clearly higher, but not a catastrophe.

Figs. 7A and 7B show the same threshold data points as Figs 6C,D, but now plotted in relation to the R_{LUSTRE} surface. Data points for condition 6 fell very close to the threshold contours representing a decrease in lustre, and did so consistently at the 5 pedestal levels for which reliable data were obtained. At the lowest pedestal contrast (1.8%) we did not obtain reliable discrimination thresholds across the three observers, but this also agreed with the model since the lustre response at this low contrast was too weak to generate a threshold contour. On the other hand, conditions 3,4,5 (already well explained by R_{MAX}) generally did *not* fall close to the threshold contours for lustre.

Lastly, Figs. 7C,D show the lustre analysis for condition 8 (*dichoptic antiphase*). Here the monocular pedestal points lie on the c_L or c_R axis, where lustre is zero, and increasing ΔC *increases* the model's lustre response. The threshold contour for condition 8 (yellow) is defined by $R_{LUSTRE} = \sigma$, and is necessarily the same for all pedestal contrasts. Data for the three higher pedestal contrasts (Fig. 7C) fell very close to this contour, as did the data for two of the three lower contrast pedestals shown in Fig. 7D.

In summary, most of the antiphase data are accurately accounted for by the contrast cue, R_{MAX}, and the remainder are well explained by the lustre cue. Fig. 8 helps to clarify and quantify this key point. Here we computed performance and discrimination thresholds based on the two cues separately. In conditions 3 and 5, thresholds from lustre (dashed curves) were far too high at all contrasts, but thresholds from R_{MAX} (solid, coloured curves) closely matched the data. The reverse was true for condition 6, where, as discussed above (Fig. 6C,D), no threshold could be measured for R_{MAX}, but the data closely matched the thresholds from lustre. In general lustre was too weak to be useful at low pedestal contrasts but, in conditions 8 and 4, cue use depended on contrast level. In condition 8, lustre was the more effective cue at higher pedestal contrasts, but R_{MAX} was the only useful cue at low contrasts, below 3%. In condition 4, both cues were useful at higher contrasts. Conditions 8 and 4 (Fig. 8) illustrate how efficient use of the contrast and lustre cues together (thick grey curves) provides a more precise account of the results than either cue alone.

5 Discussion

We measured the human visual system's ability to distinguish changes in contrast for eleven different monocular, binocular and dichoptic (antiphase) conditions, across a wide range of contrast levels. To our knowledge, this is the most comprehensive study of contrast discrimination to date, and it provides a stringent test for models of binocular contrast processing. We found that the most useful way to visualize the tasks, the model responses and experimental data was in *binocular contrast space* (Figs. 2, 5, 6, 7). The response of any mechanism can be rendered as a 3D surface in this space (Fig. 5), and if visual performance depends mainly on that mechanism then discrimination thesholds for a given pedestal should fall on a specific pair of iso-response contours on that surface. For binocular *in-phase* pedestals that was found to be correct: a single response surface (R_{MAX}) captured all the data points very well (Fig. 6A,B). The clusters of data gave direct information about the surface shape in quite a large neighbourhood around each pedestal point.

5.1 Contrast cue from monocular & binocular channels

In our model, the R_{MAX} surface arises as the response envelope (*max*) over six input mechanisms - the left-eye, right-eye, and binocular channels for positive contrast and for negative contrast (Fig. 5, stage 2). This model incorporates the binocular channel that we proposed previously (Meese et al, 2006), and extends it by adding the parallel monocular channels. This extension did not add any free parameters. It is supported by our finding that when the model was re-fitted without the monocular channels (their responses were set to 0) the fit was poor for conditions 9 and 11, but good for all other conditions including antiphase pedestals (Fig. S7). Success for the antiphase conditions, *without* the Mon channels, rests on (i) the existence of separate channels for the two polarities, so that out-of-phase cancellation does not occur in the binocular responses, and (ii) the fact that Mon and Bin channel responses are the same for antiphase conditions (Fig. 5), so that removing Mon channels had no effect on model responses in

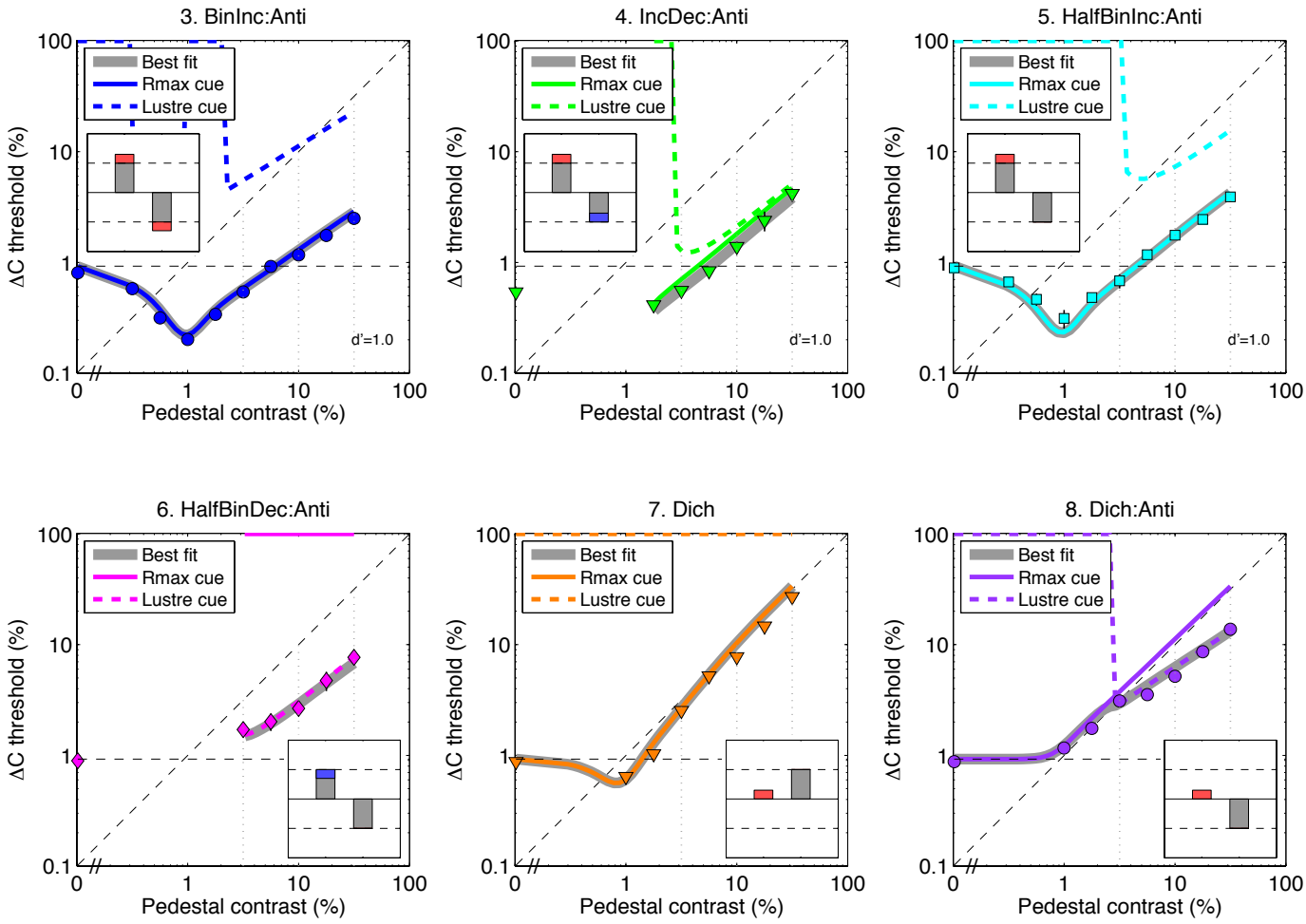

Figure 8. Summary (1). How the two cues (R_{MAX}, R_{LUSTRE}) contributed to discrimination performance in the antiphase conditions. Data re-plotted from Fig. 4, conditions 3-8. Using the best-fit parameters (Table 2), model thresholds were derived using only R_{MAX} as the decision variable (coloured solid curves) or only R_{LUSTRE} (coloured dashed curves). Neither cue alone could explain performance overall, but performance was very well explained when we assumed that the observer could use both cues (full model; thick grey curves). Lustre was markedly the better cue for condition 8, and the only useful cue in condition 6, but contrast (R_{MAX}) was much the better cue in conditions 3 and 5 and of variable benefit in condition 4. The lustre cue is absent when contrast polarity is the same in both eyes; hence contrast (R_{MAX}) was the only available cue for condition 7 here, and for conditions 1,2,9,10-13 (Fig. 4).

the antiphase quadrants. On the other hand, without the Mon channels, in-phase thresholds for condition 9 were predicted to be 2 to 3 times higher than observed, and those for condition 11 were up to a factor of two lower than observed. In addition, thresholds for half-binocular increments and decrements (conditions 10 and 11) were predicted to be the same (Fig. S7), quite unlike the data where the decremental thresholds were 2-3 times higher than the incremental. These data (conditions 9,10,11) were very well fitted, however, when the monocular channels were included (Fig. 4E). Those mechanisms correctly influenced the shape of the response surface in regions away from the positive diagonal (Fig. 6A,B), not examined in any earlier studies. Thus *the need for a monocular contribution was revealed most directly by the in-phase pedestals with monocular contrast decrements* (conditions 9 and 11).

Conversely, when the model was re-fitted without the binocular channels the resulting fit was generally very poor and unsatisfactory (RMS error was 3.81 dB, three times larger than the best-fitting model). In this model framework, then, both the monocular and binocular channels made essential contributions to performance. But unlike contrast and lustre, they did not act as separate or independent cues. Instead it was their highly nonlinear interaction (the *max*) that accounted for observed performance.

Having established that both are necessary, we can now ask: what contribution do the monocular and binocular channels make to performance of the full model? We took the best-fitting full model and simply switched off (set to 0) the responses of monocular or binocular channels at stage 2 to find out what impact this had on predicted thresholds for

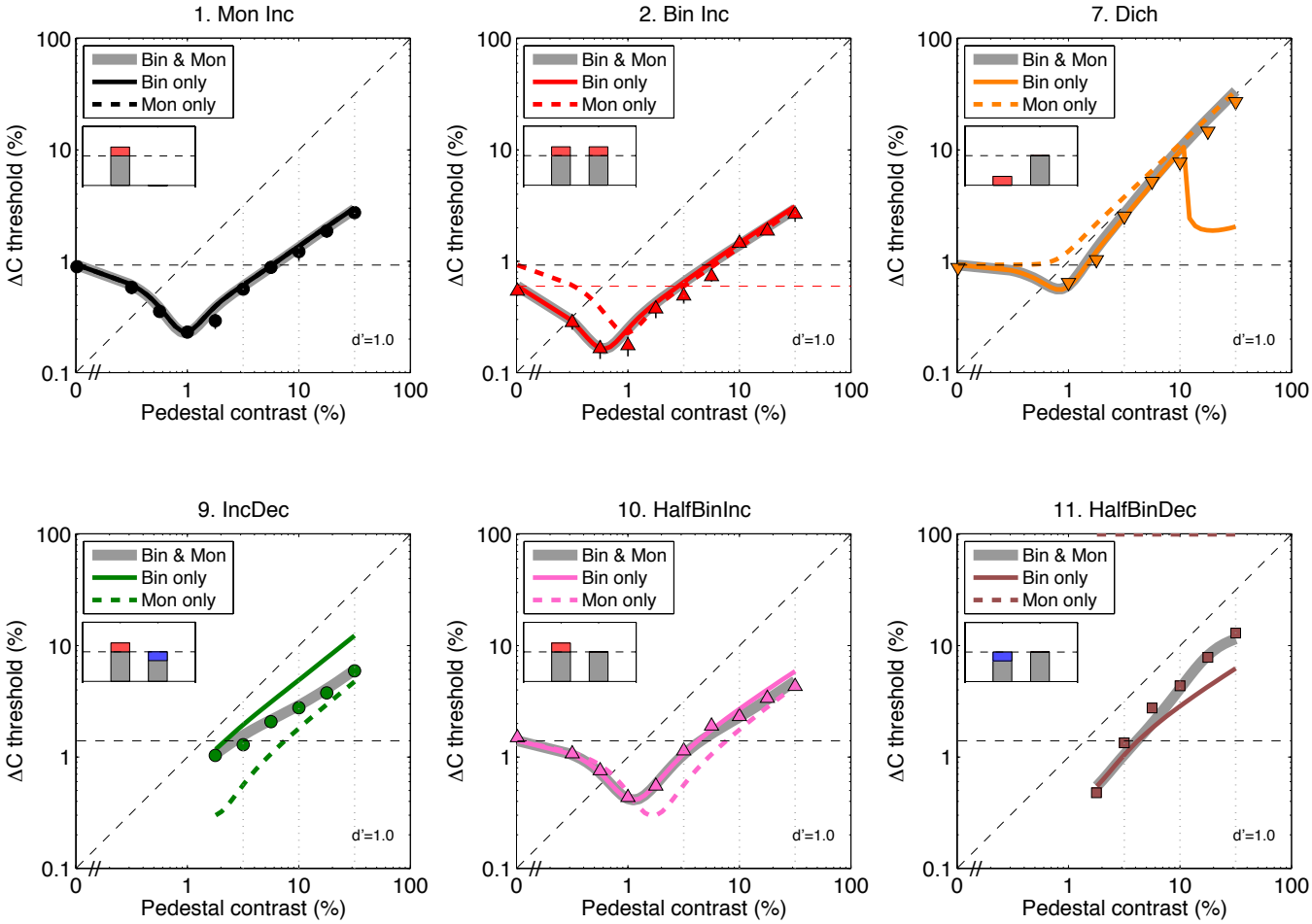

Figure 9. Summary (2). How the monocular and binocular responses contributed to discrimination performance in the in-phase conditions. Data re-plotted from Fig. 4, conditions 1,2,7,9-11. Using the best-fit parameters (Table 2), model thresholds were derived using only binocular channels (coloured solid curves) or monocular channels (coloured dashed curves). Performance overall was well explained only when both kinds of channel contributed to the contrast cue (R_{MAX}) (thick grey curves). Their contributions occurred in different regions of binocular contrast space (Fig. S5). Unlike contrast and lustre (Fig. 8) they did not act as separate cues.

in-phase conditions (Fig. 9). For example, deleting the monocular channels left predictions unchanged in conditions 1 and 2 (Fig. 9; Mon Inc, Bin Inc), so we infer that the binocular channels were sufficient for those two conditions, and largely so for condition 10 (HalfBinInc) as well. On the other hand, the monocular channels contributed to the good fit for conditions 7,9,11, because without them the 'Bin only' predictions deviated from the data. Similarly, the binocular channels were necessary in conditions 2,7,9,10,11 where the 'Mon only' predictions were insufficient (Fig. 9). Performance overall was well explained only when both kinds of channel contributed to the contrast cue (R_{MAX}) (thick grey curves).

Interestingly, Fig. 9 shows us that in some cases observers would have done better to use the monocular channels alone (conditions 9 & 10, low-medium contrasts), or the binocular channels alone (conditions 7 & 11, high contrasts). The *max* operator explains the observers' inability to do this: observers cannot freely select the most useful of these channels, and only have access to the highest value amongst them. This implies that the *max* operation is a hard-wired or obligatory process.

5.2 Contrast-matching

If we are correct that R_{MAX} represents contrast, and we have inferred its response surface shape correctly, then we can make a strong prediction: the contours of the R_{MAX} surface should not only predict contrast discrimination, but should also predict results on dichoptic contrast-matching (Baker et al., 2007, 2012). Fig. 10 shows that this prediction is accurately upheld. Data points of a given colour represent pairs of dichoptic contrasts that all match the same fixed binocular standard contrast (Fig. 10A) or monocular standard (Fig. 10B). Since they all produce the same response level, each set of points must lie on the same iso-response contour. Thick black contours (Fig. 10A) show that data from Baker et al (2012) fell very close to the model contours for all four standard contrast levels, with no free parameters. Fig. 10B shows similar contrast-matching data from another laboratory (Ding et al., 2013) again falling close to the R_{MAX}

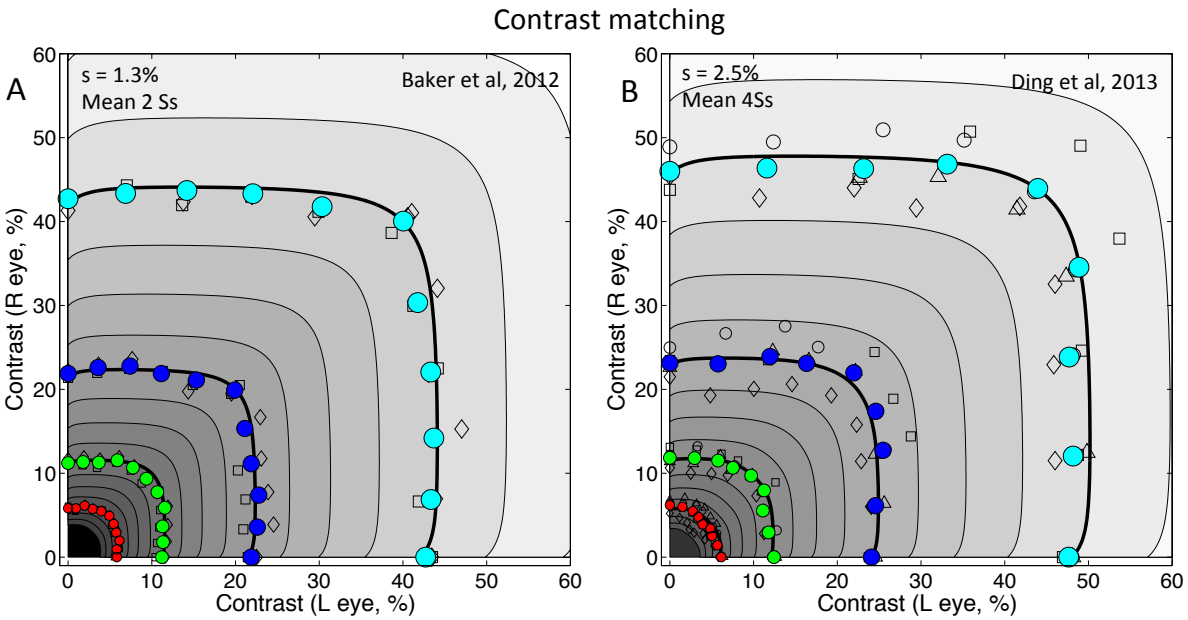

Figure 10. Contrast matching for edges (A) and gratings (B). A: Solid circles show means of 2 subjects (DHB, SAW) from Baker et al 2012, Fig. 6. These data were mirrored about the positive diagonal. Two subjects (open diamonds and squares; not mirrored) viewed test and standard images that were single, sharp, step edges, 1 deg long, shown for 200 ms with 4 standard binocular contrasts (5, 10, 20, 40%; red, green, blue, cyan). Model parameters (Table 2) were used to create the R_{MAX} response surface (grey). Black contours are the 4 iso-response contours that predict where the contrast-matching data should fall. B: Solid circles show means of 4 subjects from Ding *et al* 2013, their Figs. 9 and10. Test and standard images were sinewave gratings, with 4 standard monocular (left-eye) contrasts (6, 12, 24, 48%; red, green, blue, cyan). Two subjects (JS, KT; open diamonds and squares) viewed test durations of 117 ms; the other two (CG, CF; open circles and triangles) were tested at 1000ms. Spatial frequency = 0.68 c/deg, phase disparity = 0. Data were not mirrored about the diagonal. In B, just two minor changes were made to the model: to capture the greater linearity at low contrast (6%, red) parameter *s* was increased from 1.3% to 2.5% contrast, and to capture a slight left/right asymmetry in these mean data, we assumed a slight difference in contrast gain at the linear front-end of the model: the right eye's initial response to contrast was increased by 5% (i.e. $c_R^+ = 1.05c_R$).

surface contours (but with two minor parameter adjustments; see figure legend for details). Note how the surface contours, and the data, change shape with contrast level, showing more pronounced 'winner-take-all' behaviour at higher contrasts (replicated by Ding & Levi, 2016, their Fig. 6). These two analyses - on a task that explicitly requires judgement of contrast - provide independent confirmation that the R_{MAX} surface represents the contrast response of the binocular visual system. Note also that the data did *not* fall close to the surface contours of the binocular channel alone. That surface (Fig. 5B) exhibits a strong curvature (the Fechner paradox) that is not seen in the contrast-matching data, and which is eliminated in the model response when the monocular channels make their contribution to R^+, inherited by R_{MAX}.

Contrast-matching gives us rather precise information about the 2-D shape of the surface contours, but not about their vertical spacing or the steepness of the surface. Conversely, the discrimination tasks (Fig. 6A,B) sampled the surface contours more sparsely, but with the assumption of late additive noise (σ) they give us richer information about the 3-D surface shape and steepness. It is not trivial that both tasks are consistent with a single response surface. Hence the model surface in Fig. 10A (same surface as Fig 6A,B) unites both forms of evidence into a single picture of the binocular contrast response for horizontal, in-phase (zero disparity) image pairs.

5.3 Previous studies at detection threshold

No previous studies have tested suprathreshold discriminations with antiphase pedestals, but several have compared detection thresholds (without a pedestal) for test stimuli of the same or opposite polarity. Fig. 11 shows a quantitative comparison (see figure legend for details). Despite large differences between the stimuli and methods used, there is broad agreement across studies that stimuli of the same polarity (1st and 3rd quadrants) combine almost linearly at detection threshold, while those of opposite polarity are close to winner-take-all. Our model was fitted to our entire dataset, but it clearly fits our threshold data (red and green symbols) very well (grey curve) and also gives a good account of two earlier studies. Other models can fit these data (Cogan, 1987; Cohn & Lasley, 1976), but our proposals have the merit of being tested and supported by a much wider range of conditions than was previously available.

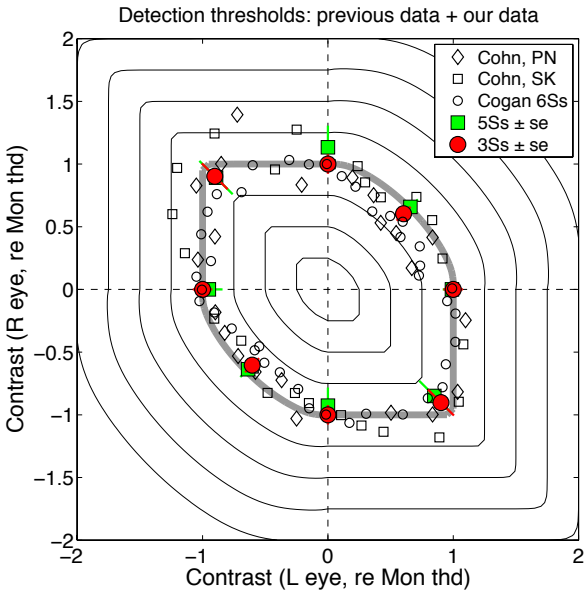

Figure 11. Dichoptic detection thresholds compared across studies. Threshold values are scaled so that the average monocular threshold = 1 for each study. Diamonds: data from Cohn & Lasley (1976), their Fig. 1, subject PN. Stimulus was a light spot, 10 min arc diameter, briefly incremented or decremented; method of adjustment (1 run). Open squares: as diamonds, but subject SK (median of 7 runs). Open circles: data from Cogan (1987), his Fig. 7; mean of 6 subjects. Stimulus was a briefly flashed increment or decrement (2 msec) of a luminous field 12 deg in diameter; method of adjustment. Red circles: data from our main experiment, with zero-contrast pedestal; means of the 3 subjects (4 independent points, 4 mirrored). Green squares: further data from our laboratory; means of 5 undergraduate subjects (no mirroring), tested in conditions very similar to the main experiment (horizontal, 1 c/deg gratings, 2AFC), except stimulus duration was 500ms. Thin curves are model R_{MAX} iso-response contours spaced at equal intervals of monocular contrast from 0 to 2%; parameters as Table 2. For the model curves, axis values represent percent contrast; thick grey curve is the iso-response contour at 1% monocular contrast.

5.4 Monocular and binocular regions of binocular contrast space

From the fitted model, we can determine *where* in binocular contrast space the monocular and binocular channels make their contribution to R_{MAX}. This is illustrated in Fig. S5 (*supplementary material*). At high contrasts (say, 20-50%), the binocular channel response dominates only in a surprisingly narrow region where the left and right contrasts are nearly equal (Fig. S5A). Outside this region, interocular suppression causes the binocular response R_B to fall below the monocular ones (R_L or R_R) and then it is the monocular channels that deliver the contrast cue via R_{MAX}. At low contrasts, however, interocular suppression is relatively weak (because the constant s in eqn. 4 is then relatively strong), and the binocular channel gains influence over a much wider range of interocular contrast ratios (Fig. S5B). As a result, binocular summation is much more directly evident at low contrasts both in the discrimination data (Fig. 4A) and detection data (Fig. 11), and in the oblique orientation of the R_{MAX} surface contours at low contrasts (Figs. 6B, 11, S5B).

5.5 Model variants

The 6-channel 2-cue model accounts for performance on a great variety of dichoptic discrimination and contrast-matching tasks with unusual accuracy. Yet, as with any model, one can ask which features of the model are necessary and which are optional. We saw above that the parallel architecture of monocular and binocular channels was not optional.

Firstly, we show that it's not crucial for the interocular suppressive terms to be polarity-specific. We re-fitted the model, putting Eqn. 1*a* in place of Eqn. 1:

$$r_L^+ = \frac{(c_L^+)^m}{s + c_L^+ + c_R^+ + c_L^- + c_R^-},\ r_L^- = \frac{(c_L^-)^m}{s + c_L^+ + c_R^+ + c_L^- + c_R^-} \qquad (1a)$$

and similarly for Eqn. 2. The suppression (denominator terms) could then arise from either polarity. This caused substantial suppression of the binocular channel response surfaces in the 2nd and 4th quadrants, but that did not carry through to later stages because it was effectively hidden by the monocular channel responses. There was little change in the $R+$, $R-$ or R_{MAX} maps, and essentially no change in the fitted parameters or the predicted pattern of thresholds. Goodness of fit (RMS error) was unchanged at 1.16dB.

Secondly, we addressed an important question about binocular summation within the binocular channel. Like-polarities sum, but do opposite polarities cancel ? At low contrasts, binocular thresholds were much lower than monocular ones (Fig. 4A), but when the gratings were out of phase, binocular thresholds were very similar to monocular ones (Fig. 4B, Fig. 11). This implies polarity-specific summation, as in Eqn. 7. To introduce the possibility of cancellation, we introduced a 'push-pull' arrangement familiar to cortical physiologists, where inputs of the non-preferred polarity carried a negative sign, rather than being ignored. This creates quasi-linear, signed summation, before half-wave rectification. Thus we replaced Eqn. 7 with Eqn. 7*a*, creating push-pull inputs (bracketed terms), which were set to 0 if negative (i.e. half-wave rectification):

$$R_B^+ = \frac{(r_L^+ + r_R^+ - r_L^- - r_R^-)^p}{z + (r_L^+ + r_R^+ - r_L^- - r_R^-)^q}, \quad R_B^- = \frac{(r_L^- + r_R^- - r_L^+ - r_R^+)^p}{z + (r_L^- + r_R^- - r_L^+ - r_R^+)^q}. \qquad (7a)$$

We then re-fitted the model and the outcome was very similar to that just described for Eqn. 1*a*. There was a dramatic change in the binocular channel response - cancellation between opposite-polarity inputs - but almost no change in later stages of response (see Fig. S9) and no change in the predicted thresholds or goodness of fit. Again the monocular responses switched in, via the *max* operator, when the binocular response fell away. The same was true when both variants (Eqns. 1*a* and 7*a*) were applied together.

In short, these two analyses show that several aspects of the binocular channel remain hidden from us. In antiphase conditions monocular channel responses play an important role, and because of this the data cannot tell us whether the binocular channel receives suppression from, or negative input from, the non-preferred polarity. The modelling reveals more clearly what we don't know, and why.

5.6 Contrast & Lustre

Our second set of key findings concerns *antiphase* pedestals. Here, despite intensive effort, we could not find a single response surface that accounted for all the observed thresholds. Instead we propose that *two* response surfaces - representing two different subjective cues, contrast and lustre - are needed to understand performance in antiphase conditions. Depending on the direction of change in binocular contrast space, some antiphase conditions (3,4,5) relied wholly or mainly on the contrast cue, while other conditions depended on lustre, responding to contrast changes that either decreased lustre (condition 6) or increased it (condition 8). The two response surfaces are rendered as contour maps in Figs. 5H, 5I. The R_{MAX} and R_{MIX} surfaces are identical when both eyes view the same contrast polarity. Hence the lustre map (R_{MIX} - R_{MAX}, Fig. 5I) is zero for in-phase gratings (1st and 3rd quadrants), but is positive in the 2nd and 4th quadrants, and peaks when the two eyes view equal and opposite contrasts (c_L=-c_R).

When the model was re-fitted with the lustre cue removed (set to zero), the fit was poor in many conditions, especially for antiphase pedestals (Fig. S6). Overall RMS error was more than doubled (to 2.6 dB), and in a nested-model comparison the fit was significantly worse than the full model [F(1,102) = 422.9, P<0.00001]. Thus, within our model framework, the lustre cue was necessary to explain performance accurately in the antiphase conditions.

In a study of binocular rivalry between opposite polarities, Whittle (1965) conjectured that "Lustre ... occurs when stimuli to both rivalry (contours of opposite sign) and fusion (contours of the same sign) are presented." Our antiphase gratings do not contain contours of the same sign in the same location, but our scheme could still satisfy Whittle's conjecture. R_{MAX} might be the code normally associated with single vision (either by fusion or by suppression; Georgeson & Wallis, 2014), while R_{LUSTRE} is the signal that encodes the simultaneous presence of opposite signs. We now consider whether any other evidence supports this view that lustre arises solely from opposite polarities, and not more generally from the ability to sense differences in contrast (Formankiewicz & Mollon, 2009).

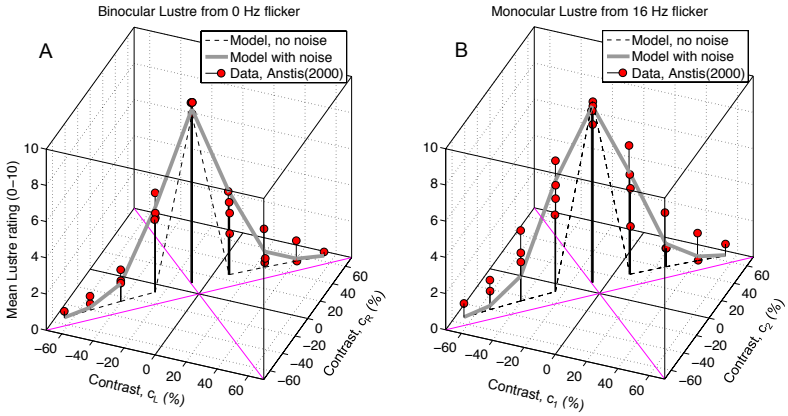

Figure 12. Ratings of lustre (symbols; from Anstis, 2000, Fig 2a, re-plotted in a new format), are compared with predictions based on R_{LUSTRE}, computed for the conditions of Anstis's experiment and scaled to the range 0-10. A: Binocular lustre. Each data point represents subjects' mean lustre rating for a dichoptic pair of achromatic spots with different luminances. We converted the spot luminances to Michelson contrast (c_L,c_R) in percent, and plotted the results in binocular contrast space. Predictions based on R_{LUSTRE} are shown without luminance noise (dashed curve), and with luminance noise (solid grey curve). B: As A, but for 'monocular lustre', where the same spot was shown to each eye, but spot contrast switched between c_1 and c_2 over time at 16 Hz. For consistency, the same model was applied to both A and B, and it assumed temporal smoothing of the R+ and R- signals before R_{LUSTRE} was computed. Temporal smoothing was a key factor at 16Hz (B), but has no effect at 0 Hz (A). See Discussion *Lustre judgements* for details.

5.7 Lustre judgements

Despite much observation and discussion, the relations between lustre, gloss, rivalry and transparency are poorly understood, and there have been few quantitative studies to define the necessary conditions for lustre. Anstis (2000) asked subjects to rate their impressions of lustre for dichoptic pairs of achromatic spots, with different luminances that were higher and/or lower than the background luminance. Several background levels were used. We converted his spot luminances to Michelson contrast (c_L,c_R), and plotted the mean lustre ratings in binocular contrast space (Fig. 12A). Lustre was highest when the spots had opposite contrast polarity, as our model predicts, but was also fairly high at adjacent points lying in the first and third quadrants, where contrasts were different, but of the same sign. R_{LUSTRE} (by design) predicted lustre (dashed curve) only when the spots had opposite contrast polarities - at just one point in the 2nd quadrant for this experiment. But when noise was added to all luminance levels (both target and background; gamma distribution, s.d. = 0.7*mean) this perturbed the contrast values, and so some same-polarity pairs became opposite-polarity, at least some of the time. Averaged over many samples, the grey curve shows that the resulting mean value of R_{LUSTRE} fitted the data very satisfactorily. We conclude that lustre is mainly induced by opposite signs of contrast, and that luminance noise can explain why lustre diffuses into the same-polarity quadrants (Fig. 12A). As Helmholtz remarked: "If one eye sees black, and the other eye sees white [in corresponding locations] the impression will be that of a surface shedding a pale lustre" (von Helmholtz, 1925, p.514).

Anstis (2000) also examined 'monocular lustre', where the same spot was shown to both eyes, but spot contrast switched between c_1 and c_2 over time at 16 Hz. The lustre ratings (Fig. 12B) showed a very similar pattern to binocular lustre. Monocular lustre can thus occur when different contrasts are alternated over time. In our model two separate signals, R+ and R-, carry information about opposite signs of contrast, and if these signals are temporally smoothed then the alternating R+ and R- signals will come to overlap in time, and so generate a lustre response. To test this idea against Anstis's data, we implemented a dynamic version of the model that assumed temporal smoothing of the R+ and R- signals (low-pass filtering; integration time about 50 ms) before R_{LUSTRE} was computed. This smoothing produced a time-varying lustre response, and lustre rating was taken to be proportional to the time-averaged value of R_{LUSTRE}. With the input luminances perturbed by noise as before, there was a good fit between model and data (Fig. 12B). This provides some direct support for our model of lustre (Fig. 3), and shows how binocular and monocular lustre can arise from the same set of mechanisms.

5.8 Limitations & future challenges

Like many previous models, the present one treats the ocular contrasts c_L, c_R as the system's input values. Luminance contrast, for this model, is a pointwise primitive quantity. We have not tried to give any account of how the retina derives these contrast values from photoreceptor responses, and we have not explicitly represented the variation of stimuli and responses across visual space. The present model therefore cannot deal with phase disparities other than 0 and 180^{O}. A more complete model would include the spatial (x,y) dimensions, and could then address other types of experiment such as the judgment of binocular spatial phase (Ding & Sperling, 2006) or the binocular fusion of edges (Georgeson & Wallis, 2014). Meese & Baker (2011) on the other hand, did include early, monocular spatial filtering in their model for binocular summation and spatial summation of contrast. They concluded that local, phase-specific, binocular summation of contrast responses precedes a second-stage of broader spatial summation that generalizes across both spatial phase and spatial position, and is followed by a third stage output nonlinearity. Their first and third stages correspond closely to the two-stage binocular channel of Meese et al (2006) that is also embedded in the present model.

Why should we need to include parallel monocular channels, when no previous study of contrast discrimination has needed them? This is a key question, and we think the answer is that we used a more comprehensive range of test directions in binocular contrast space, around each pedestal point (Fig. 2B). It was the conditions involving contrast decrements (9 and 11) - not tested in any previous studies - that revealed the need for monocular channels. We also showed that evidence from contrast matching experiments (Fig. 10) is not consistent with our binocular channel's response surface alone (Fig. 5B), but directly favours the R_{MAX} surface that emerges from *max*-like selection across the monocular and binocular channel responses.

We recognize that other front-ends to the model might be possible. A set of input equations that combined left and right eye inputs in a different way, but gave rise to the same $R+$ and $R-$ maps, and hence the same R_{MAX} and R_{MIX} maps, would be functionally equivalent to our model (Fig. 5). For example, the DSKL model 3c (Ding *et al*, 2013), an elaborated version of the Ding & Sperling (2006) model, gives a good account of dichoptic spatial phase and contrast perception. It produces the appearance of winner-take-all (WTA) behaviour between the eyes (similar to the rounded-square binocular-response contours of Fig. 6A or Fig.10),

but unlike our model it does not take the *max* over monocular and binocular channels to achieve this WTA effect. Instead the DSKL model re-shapes the left and right eye input amplitudes before binocular summation, using a combination of interocular suppression and interocular facilitation or gain enhancement. This approach readily handles dichoptic spatial phase perception, but at the cost of two extra free parameters, along with the conceptual complexity of several interacting gain-control mechanisms.

It would be parsimonious if the complexities of binocular interaction could all be handled in this way, by elaborating the mechanism of binocular summation. But antiphase gratings reveal a limit to this approach. A core assumption (Ding & Sperling, 2006; Ding & Levi, 2016) has been that only the summed output is available (Fig. 1a). On this view, binocular interactions of various kinds modify the monocular signal amplitudes, and these modified sine-wave signals are then summed arithmetically. But if gratings of equal contrast are presented to the two eyes then whatever forces shape the left- and right-eye amplitudes they must by symmetry remain equal. And those equal-and-opposite sinewaves must then completely cancel each other in the sum. Hence antiphase gratings should be invisible, but they are not. Detection thresholds (Fig. 11) and contrast-matching data imply that we see one or other of the monocular contrasts (Baker et al., 2012). An additional mechanism seems inevitable.

We have shown here that having separate mechanisms for opposite polarities prevents such cancellation and, via the contrast and lustre cues, enables a good account of antiphase contrast discrimination. This was true whether the monocular channels were explicitly included (Fig. 4) or not (Fig. S7). Nevertheless, the monocular channels in our model were essential: they created the winner-take-all effect that was vital for understanding both contrast discrimination and matching for dichoptic in-phase gratings. And in a variant of our model where antiphase cancellation was included (Eqn. 7a, above), the monocular channels became essential for antiphase conditions as well.

Future work, however, should examine another interesting possibility - that binocular difference channels (Cohn & Lasley, 1976; Cohn, Leong, & Lasley, 1981; Jennings & Kingdom, 2016; May, Zhaoping, & Hibbard, 2012) play a role in these discrimination tasks. In one sense, our model already contains a 'difference channel', because R_{LUSTRE} is a response to the presence of opposite contrasts in the two eyes, but it does not respond more generally to a simple contrast difference where the polarity is the same in both eyes. Another possibility is that the monocular channels in our model might be replaced by opponent channels that compute ocular contrast difference (L-R, and R-L). This has yet to be explored.

The experiments here are spatially one-dimensional; they do not consider 2-D interaction effects such as cross-orientation suppression. Previous experiments have shown that cross-orientation suppression takes place both within and between the eyes, placing those monocular and dichoptic suppressive influences at stage 1 of the two-stage model (Baker, Meese & Summers, 2007; Meese & Baker, 2009). This is readily accommodated by stage 1 of the binocular pathway here, though whether the purely monocular channels exhibit cross-orientation suppression also remains to be explored.

The present model has no spatial dimension, and we expect that extending it from a model of contrast coding to a model of binocular spatial vision will lead to new and interesting theoretical developments (cf. Ding & Levi, 2016).

6 Conclusions

The 6-channel 2-cue model described here (Fig. 3) accounts very well for eleven forms of binocular contrast discrimination function. The model subsumes our earlier one (Meese et al, 2006) that had binocular summation and interocular suppression but no monocular channels in parallel with the binocular ones. With only two extra free parameters (n,a) it explains several key effects where the earlier model failed. It is parsimonious because contrast gain parameters (m,s,z,p,q) are the same for all channels. Lustre emerged as an important additional cue in some, but not all, antiphase discrimination tasks.

In brief, the theoretical questions we posed about contrast coding in binocular vision, and the model-based answers we propose, are these:

1. Do we have separate, parallel, monocular & binocular pathways? *Yes, up to a point*
2. Do we have separate pathways for opposite contrast polarities? *Yes*
3. Do the monocular pathways have suppressive interactions between eyes? *No*
4. Do the binocular pathways have suppressive interactions between eyes? *Yes*
5. Is binocular summation polarity-specific? *Yes*
6. Does binocular summation entail cancellation between opposite polarities? *Can't tell*
7. Is interocular suppression polarity-specific? *Can't tell*
8. Do we have independent perceptual access to these early pathways? *No, only to the max*
9. How many perceptual outputs or cues are used in these tasks? *Two: contrast and lustre*
10. Do we have independent perceptual access to these two cues? *Yes*

Acknowledgments

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC grant BBH00159X1) and by the Engineering and Physical Sciences Research Council (EPSRC grant EP/H000038/1). We thank Jian Ding for supplying data re-plotted in Fig. 10B, and Stuart Anstis for data re-plotted in Fig. 12.

Data access

To access the research data underlying this publication, please see: http://dx.doi.org/10.17036/3c5e1cfc-5988-4380-843e-dadbfc093e44

References

Anstis, S. M. (2000). Monocular lustre from flicker. *Vision Research*, *40*, 2551–2556. doi:10.1016/S0042-6989(00)00131-0

Anstis, S. M., & Duncan, K. (1983). Separate motion aftereffects from each eye and from both eyes. *Vision Research*, *23*(2), 161–169. doi:10.1016/0042-6989(83)90139-6

Anzai, A., Bearse, M. A., Freeman, R. D., & Cai, D. (1995). Contrast coding by cells in the cat's striate cortex: monocular vs. binocular detection. *Visual Neuroscience*, *12*, 77–93. doi:10.1017/S0952523800007331

Baker, D. H., & Meese, T. S. (2007). Binocular contrast interactions: dichoptic masking is not a single process. *Vision Research*, *47*(24), 3096–107. doi:10.1016/j.visres.2007.08.013

Baker, D. H., Meese, T. S., & Georgeson, M. A. (2007). Binocular interaction: contrast matching and contrast discrimination are predicted by the same model. *Spatial Vision*, *20*(5), 397–413. doi:10.1163/156856807781503622

Baker, D. H., Meese, T. S., & Hess, R. F. (2008). Contrast masking in strabismic amblyopia: attenuation, noise, interocular suppression and binocular summation. *Vision Research*, *48*(15), 1625–40. doi:10.1016/j.visres.2008.04.017

Baker, D. H., Meese, T. S., & Summers, R. J. (2007). Psychophysical evidence for two routes to suppression before binocular summation of signals in human vision. *Neuroscience*, *146*(1), 435–48. doi:10.1016/j.neuroscience.2007.01.030

Baker, D. H., Wallis, S. A., Georgeson, M. A., & Meese, T. S. (2012). The effect of interocular phase difference on perceived contrast. *PLoS ONE*, *7*(4), 1–6. doi:10.1371/journal.pone.0034696

Bearse, M. A., & Freeman, R. D. (1994). Binocular summation in orientation discrimination depends on stimulus contrast and duration. *Vision Research*, *34*(1), 19–29. doi:10.1016/0042-6989(94)90253-4

Bixby, F. L. (1928). A phenomenological study of luster. *Journal of General Psychology*, *1*(1), 136–174.

Blake, R., & Fox, R. (1973). The psychophysical inquiry into binocular summation. *Perception & Psychophysics*, *14*(1), 161–185.

Blake, R., Overton, R., & Lema-Stern, S. (1981). Interocular transfer of visual aftereffects. *Journal of Experimental Psychology. Human Perception and Performance*, *7*(2), 367–81.

Blake, R., Sloane, M., & Fox, R. (1981). Further developments in binocular summation. *Perception & Psychophysics*, *30*(3), 266–76.

Campbell, F. W., & Green, D. G. (1965). Monocular versus binocular visual acuity. *Nature*, *208*, 191–192.

Cogan, A. I. (1987). Human binocular interaction: towards a neural model. *Vision Research*, *27*(12), 2125–2139. doi:10.1016/0042-6989(87)90127-1

Cohn, T. E., & Lasley, D. J. (1976). Binocular vision: two possible central interactions between signals from two eyes. *Science*, *192*(4239), 561–563. doi:10.1126/science.1257791

Cohn, T. E., Leong, H., & Lasley, D. J. (1981). Binocular luminance detection: availability of more than one central interaction. *Vision Research*, *21*, 1017–1023.

Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, *27*(11), 1367–1377. doi:10.1068/p271367

Ding, J., Klein, S. A., & Levi, D. M. (2013). Binocular combination of phase and contrast explained by a gain-control and gain-enhancement model. *Journal of Vision*, *13(2):13*, 1–37. doi:10.1167/13.2.13.doi

Ding, J., & Levi, D. M. (2016). Binocular contrast discrimination needs monocular multiplicative noise. *Journal of Vision*, *16(5):12*, 1–21. doi:10.1167/16.5.12

Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(4), 1141–6. doi:10.1073/pnas.0509629103

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A*, *11*(6), 1710–1719. doi:10.1364/JOSAA.11.001710

Formankiewicz, M. A., & Mollon, J. D. (2009). The psychophysics of detecting binocular discrepancies of luminance. *Vision Research*, *49*(15), 1929–1938. doi:10.1016/j.visres.2009.05.001

Gawne, T. J., & Martin, J. M. (2002). Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. *Journal of Neurophysiology*, *88*(3), 1128–1135. doi:10.1152/jn.00131.2002

Georgeson, M. A., & Wallis, S. A. (2014). Binocular fusion, suppression and diplopia for blurred edges. *Ophthalmic and Physiological Optics*, *34*, 163–185. doi:10.1111/opo.12108

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Green, M., & Blake, R. (1981). Phase effects in monoptic and dichoptic temporal integration: flicker and motion detection. *Vision Research*, *21*, 365–372.

Haynes, J.-D., Deichmann, R., & Rees, G. (2005). Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus. *Nature*, *438*(7067), 496–9. doi:10.1038/nature04169

Home, R. (1978). Binocular summation: a study of contrast sensitivity, visual acuity and recognition. *Vision Research*, *18*(5), 579–585. doi:10.1016/0042-6989(78)90206-7

Hou, F., Huang, C.-B., Liang, J., Zhou, Y., & Lu, Z. (2013). Contrast gain-control in stereo depth and cyclopean contrast perception. *Journal of Vision*, *13(8):3*, 1–19. doi:10.1167/13.8.3.doi

Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford: Oxford University Press.

Huang, C.-B., Zhou, J., Zhou, Y., & Lu, Z.-L. (2010). Contrast and phase combination in binocular vision. *PloS One*, *5*(12), e15075. doi:10.1371/journal.pone.0015075

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*, 106–154.

Jennings, B. J., & Kingdom, F. A. A. (2016). Detection of between-eye differences in color: Interactions with luminance. *Journal of Vision*, *16(3):23*, 1–12. doi:10.1167/16.3.23

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press.

Lampl, I., Ferster, D., Poggio, T., & Riesenhuber, M. (2004). Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex. *Journal of Neurophysiology*, *92*(5), 2704–2713. doi:10.1152/jn.00060.2004

Legge, G. E. (1979). Spatial frequency masking in human vision: binocular interactions. *Journal of the Optical Society of America*, *69*(6), 838–847.

Legge, G. E. (1984). Binocular contrast summation II. Quadratic summation. *Vision Research*, *24*(4), 385–394.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision, *Journal of the Optical Society of America*, *70*(12), 1458–1471.

Legge, G. E., & Rubin, G. S. (1981). Binocular interactions in suprathreshold contrast perception. *Perception & Psychophysics*, *30*(1), 49–61. doi:10.3758/BF03206136

Maehara, G., & Goryo, K. (2005). Binocular, monocular and dichoptic pattern masking. *Optical Review*, *12*(2), 76–82. doi:10.1007/s10043-004-0076-5

Mausfeld, R., Wendt, G., & Golz, J. (2014). Lustrous material appearances: Internal and external constraints on triggering conditions for binocular lustre. *I-Perception*, *5*, 1–19. doi:10.1068/i0603

May, K. A., Zhaoping, L., & Hibbard, P. B. (2012). Perceived direction of motion determined by adaptation to static binocular images. *Current Biology*, *22*(1), 28–32. doi:10.1016/j.cub.2011.11.025

Meese, T. S., & Baker, D. H. (2009). Cross-orientation masking is speed invariant between ocular pathways but speed dependent within them. *Journal of Vision*, 9(5):2, 1–15, http://journalofvision.org/9/5/2/, doi:10.1167/9.5.2.

Meese, T. S., & Baker, D. H. (2011). Contrast summation across eyes and space is revealed along the entire dipper function by a " Swiss cheese " stimulus. *Journal of Vision*, *11(1):23*, 1–23. doi:10.1167/11.1.23

Meese, T. S., Georgeson, M. A., & Baker, D. H. (2006). Binocular contrast vision at and above threshold. *Journal of Vision*, *6*(11), 1224–43. doi:10.1167/6.11.7

Meese, T. S., & Hess, R. F. (2004). Low spatial frequencies are suppressively masked across spatial scale, orientation, field position, and eye of origin. *Journal of Vision*, *4*, 843–859. doi:10.1167/4.10.

Moradi, F., & Heeger, D. J. (2009). Inter-ocular contrast normalization in human visual cortex. *Journal of Vision*, *9*(3):13, 1–22. doi:10.1167/9.3.13

Moulden, B. (1980). After-effects and the integration of patterns of neural activity within a channel. *Philosophical Transactions of the Royal Society B*, *290*, 39–55.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025.

Schor, C. M., Wood, I., & Ogawa, J. (1984). Binocular sensory fusion is limited by spatial resolution. *Vision Research*, *24*(7), 661–665.

Simmons, D. R. (2005). The binocular combination of chromatic contrast. *Perception*, *34*, 1035–1042. doi:10.1068/p5279

Simmons, D. R., & Kingdom, F. A. A. (1998). On the binocular summation of chromatic contrast. *Vision Research*, *38*(8), 1063–1071. doi:10.1016/S0042-6989(97)00272-1

von Helmholtz, H. (1925). *Treatise on Physiological Optics, Vol. III*. (J. P. C. Southall, Ed.). Optical Society of America.

Westendorf, D. H., & Fox, R. (1973). Binocular detection of positive and negative flashes. *Perception & Psychophysics*, *15*(1), 61–65.

Whittle, P. (1965). Binocular rivalry and the contrast at contours. *The Quarterly Journal of Experimental Psychology*, 17(3), 217–226. http://doi.org/10.1080/17470216508416435

Wolfe, J. M., & Franzel, S. L. (1988). Binocularity and visual search. *Perception & Psychophysics*, *44*(1), 81–93. doi:10.3758/BF03207480

Yu, A. J., Giese, M. A., & Poggio, T. A. (2002). Biophysiologically plausible implementations of the maximum operation. *Neural Computation*, *14*(12), 2857–2881. doi:10.1162/089976602760805313

Zhou, J., Georgeson, M. A., & Hess, R. F. (2014). Linear binocular combination of responses to contrast modulation: Contrast-weighted summation in first- and second-order vision. *Journal of Vision*, *14(13):24*, 1–19. doi:10.1167/14.13.24.doi

Appendix 1

Model equations

These equations define the model fitted to the data in Figures 4-9. Possible variations on this model are considered in the Discussion section *Model variants*.

Input: Polarity-specificity
We first separate each eye's contrast value into two sign-specific parts, both of which are non-negative:

$$c_L^+ = |c_L| \text{ if } c_L > 0, \text{ else } 0;$$
$$c_L^- = |c_L| \text{ if } c_L < 0, \text{ else } 0;$$

and similarly for the right eye.

Stage 1
• The first stage of the polarity-specific binocular channels is driven by the contrast in one eye, but has divisive contrast gain control from the same polarity in both eyes:

$$r_L^+ = \frac{(c_L^+)^m}{s + c_L^+ + c_R^+},\ r_L^- = \frac{(c_L^-)^m}{s + c_L^- + c_R^-} \qquad (1)$$

$$r_R^+ = \frac{(c_R^+)^m}{s + c_L^+ + c_R^+},\ r_R^- = \frac{(c_R^-)^m}{s + c_L^- + c_R^-} \qquad (2).$$

Note that each of these responses can be de-composed into two parts, an ocular weighting term driven by relative contrast in the two eyes, coupled with a compression of the input contrast, where all contrast terms refer only to the preferred polarity. Thus we can re-write r_L^+ as:

$$r_L^+ = w_L^+ (c_L^+)^{m-1} \qquad (3)$$

where

$$w_L^+ = \frac{c_L^+}{s + c_L^+ + c_R^+} \qquad (4)$$

and similarly for the other 3 expressions of Eqns 1, 2.

• The first stage of polarity-specific monocular channels for Left and Right eyes is the same as equations 1, 2 above, except that interocular suppression is deleted:

$$u_L^+ = \frac{(c_L^+)^m}{s + c_L^+},\ u_L^- = \frac{(c_L^-)^m}{s + c_L^-}, \qquad (5)$$

$$u_R^+ = \frac{(c_R^+)^m}{s + c_R^+},\ u_R^- = \frac{(c_R^-)^m}{s + c_R^-} \qquad (6).$$

Stage 2

• The second stage of the binocular channel sums like-polarity responses from the two eyes (e.g. r_L^+, r_R^+), and the sum is subjected to a response nonlinearity (Legge & Foley, 1980), which acts like a smooth threshold at low response levels and a power law transformation with exponent (*p-q*) at high response levels:

$$R_B^+ = \frac{(r_L^+ + r_R^+)^p}{z + (r_L^+ + r_R^+)^q},\ R_B^- = \frac{(r_L^- + r_R^-)^p}{z + (r_L^- + r_R^-)^q}. \qquad (7)$$

From Eqns. 3 and 7, we note that power-law transformations at stages 1 and 2 are in series. In simple monocular or binocular viewing this is equivalent to a single power-law whose exponent is the product of the two exponents *(m-1)(p-q)*.

• The second stage of the monocular channels is like Eqn. 7, but opposite-eye terms are again deleted:

$$R_L^+ = \frac{(u_L^+)^p}{z + (u_L^+)^q},\ R_L^- = \frac{(u_L^-)^p}{z + (u_L^-)^q}, \qquad (8)$$

$$R_R^+ = \frac{(u_R^+)^p}{z + (u_R^+)^q},\ R_R^- = \frac{(u_R^-)^p}{z + (u_R^-)^q}. \qquad (9)$$

Signal selection

MAX operators play an important role in hierarchical models of visual object recognition (Riesenhuber & Poggio, 1999), and their implementation in cortical circuits of V1 and V4 is an active area of theoretical and physiological research (Gawne & Martin, 2002; Lampl, Ferster, Poggio, & Riesenhuber, 2004; Yu, Giese, & Poggio, 2002). As introduced above, we resolved the three responses (subscripted *L,R,B*) into one, via a *MAX*-like operator, and did this separately for each polarity to create just two polarity-specific responses. We implemented the *MAX* operation via a Minkowski sum (power sum) with a high exponent *n*. Thus the two responses are:

$$R^+ = \left\{ \sum_{j=L,R,B} (R_j^+)^n \right\}^{1/n}, \qquad R^- = \left\{ \sum_{j=L,R,B} (R_j^-)^n \right\}^{1/n} \quad (10)$$

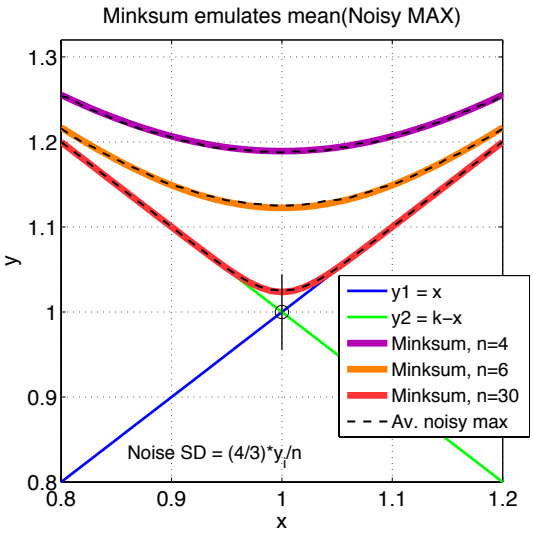

Figure A1. Minkowski sum emulates the mean output of a noisy *MAX* operator. We consider combining two variables (envisaged as neural responses) y_1, y_2 (blue, green). At each point *x*, we combine y_1, y_2 in two ways to produce an output *y*: (i) as a Minkowski sum with no noise, $y = (y_1^n + y_2^n)^{1/n}$ (coloured curves, where exponent *n*=4,6,30), and (ii) as the output of a *MAX* operator, $y = max(y_1 + \epsilon_1, y_2 + \epsilon_2)$, averaged over many independent samples ϵ_1, ϵ_2 of zero-mean Gaussian noise $N(0,\sigma_i)$. Black dashed curves show that, provided *n*>2, this mean output of the *MAX* operator closely matches the Minkowski sum when the noise standard deviation σ_i is proportional to the mean input y_i and inversely related to the Minkowski exponent *n*: $\sigma_i = (4/3)y_i/n, (i = 1,2)$. Error bars at y_1=y_2=1 show ±1 s.d. of the noise for *n*=30. Lower exponents in the Minkowski sum correspond to higher noise levels in the *MAX* operator, and in both cases the combined output value is higher than the simple noise-free maximum, $max(y_1, y_2)$. Deviation of *y* from the simple *max* is greatest when $y_1 = y_2$, and at this point $y = 2^{1/n} y_1$. Equivalently, we may say that for two equal signals the summation gain factor is $2^{1/n}$. Conclusion: when a model uses the Minkowski sum as a formalism for combining signals, it could be interpreted equally well as the nonlinear sum (power sum) of the signals, or as the mean output of a *MAX* operator with noisy inputs.

We found that the Minkowski sum (with n as a free parameter in the model fitting) gave much better fits than a simple max operation. We show in Fig. A1 that there is an interesting and perhaps unsuspected relation between the Minkowski sum and the max operator. The Minkowski sum (with no noise) is almost exactly equal to the mean output of a true max operator where each of the input signals is noisy (Fig. A1). Higher noise tends to raise the mean output in a way that is equivalent to a lower Minkowski exponent n. Thus the response R+ or R- in eqn. 10 may be interpreted as the average of the max of 3 noisy input signals. This average is slightly higher than the max of the 3 inputs without noise and this feature seems to be important in capturing the observed human performance. Fig. S1 (Supplementary Material) gives some further insight into this.

Perceptual cues: 1. Contrast

To derive a code for contrast from the six stage 2 outputs, we take the *max* (see rationale above) and this can be implemented as a second Minkowski sum with exponent *n*, taken over R^+ and R^- (see Fig. 3):

$$R_{MAX} = \left\{ (R^+)^n + (R^-)^n \right\}^{1/n} \qquad (11),$$

where *n* is expected to be large (eg. *n*>20). This single number, R_{MAX}, is the model's internal representation or code for luminance contrast.

Perceptual cues: 2. Lustre

Following our rationale and formulation of lustre (above) we used a general Minkowski sum for pooling the two polarity-specific responses:

$$R_{MIX} = \left\{ (R^+)^a + (R^-)^a \right\}^{1/a} \qquad (12),$$

where *a* is a free parameter, expected to be relatively small (eg. *a*<5) to give more substantive pooling than the *MAX* operator. We then defined the response to lustre as:

$$R_{LUSTRE} = R_{MIX} - R_{MAX} \qquad (13).$$

Eqns. 12,13 satisfy our requirement for selectivity of R_{LUSTRE}. For a non-lustrous input (ie. opposite polarities not present), either $R^+ = 0$ or $R^- = 0$, and

from Eqns. 11,12 this implies $R_{MIX} = R_{MAX}$, hence $R_{LUSTRE} = 0$, as required.

Decision processes: sensory cues, noise & observer strategies

Discrimination performance (d') for a given (*test, pedestal*) pairing can be defined separately for the two cues:

$$d'_{CONTRAST} = abs[\, R_{MAX}(test + ped) - R_{MAX}(ped)]/\sigma \quad (14),$$

$$d'_{LUSTRE} = abs[\, R_{LUSTRE}(test + ped) - R_{LUSTRE}(ped)]/\sigma \quad (15).$$

This is a late-noise model, in which both cues are perturbed by additive Gaussian noise. Observed performance (d'_{OBS}) depends on how efficiently the observer can make use of the two cues. We assumed that the two cues were independently noisy but with the same noise variance (σ^2), and that the observer could make good use of both cues:

$$d'_{OBS} = \sqrt{{d'_{CONTRAST}}^2 + {d'_{LUSTRE}}^2} \quad (16).$$

This quadratic sum represents optimal use of the cues (Green & Swets, 1966). It is an important, parameter-free, benchmark but was not a crucial assumption for our dataset. In practice, we found almost the same model performance with sub-optimal cue combination, represented by a Minkowski sum of d' values with higher exponents, from 2 (optimal) to 200 (representing the *max* of the two d' values). The exponent value is most crucial when the two d' values are very similar and that occurred only in condition 4.

Fitting the model

The equations fully define the model, and allow us to compute model performance (d'_{OBS}) for any specified test condition with pedestal contrast C and contrast change ΔC. For each C, d' was computed for a wide range of values of ΔC at 1.3 dB intervals, and the threshold value of ΔC, where d'=1, was found by interpolation. Root-mean-square (RMS) error between model and observed thresholds was computed in dB, and model parameters were adjusted by the *Simplex* algorithm (*fminsearch* in *Matlab*) to find the lowest RMS error. Multiple fitting runs (usually 20) were done with starting values randomly jittered around a plausible set of initial values, to ensure that the best-fit did not represent a local minimum in the error surface. There were in principle 8 free parameters [n,m,s,p,q,z,σ,a]. Six of these had the same meaning as in our earlier binocular-channel model (Meese et al, 2006), while two new ones (*n, a*) defined the form of pooling in the *max* and *mix* operators (Eqns. 11, 12). A ninth free parameter (z_2) was a pragmatic addition. Conditions 9-13 were drawn from our study in which stimulus duration was 100ms rather than 200ms (see Table 1) and, perhaps for this reason, contrast thresholds in the low pedestal-contrast region (below 1%) were about 3dB higher than observed for conditions 1-8. We found that just one change - allowing a higher value (z_2) in place of z for conditions 9-13 - was a simple and sufficient compensation for the procedural differences between the three studies that formed our dataset. Great explanatory power, including the power to reject unsatisfactory explanations, was gained by requiring the model to fit data from so many different conditions (total N=111 data points) simultaneously.

Supplementary Material

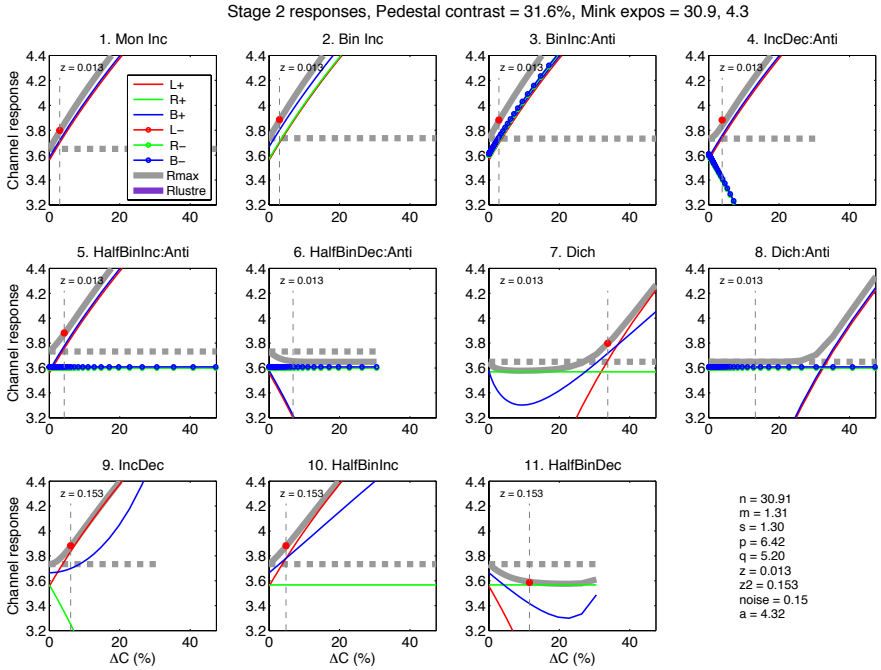

Figure S1. Diagnostic diagrams for the fitted model reveal which signals are important in which conditions. Each panel shows the six stage-2 responses (thin curves, red, green, blue; see legend), as a function of ΔC for a given task (conditions 1-11) at the highest pedestal contrast (32%). Some responses may be low or zero, hence off the scale. The 6 channel responses are accessible to perception only via the *MIX* and *MAX* operators (Fig. 3) that lead to outputs R_{MAX} and R_{LUSTRE}. Thick grey curve shows the output R_{MAX}, while the thick purple curve that would show R_{LUSTRE} is well below the plotted range here (but shown in Figure S2). The dashed horizontal grey lines show the R_{MAX} output level for the pedestal-only interval (ΔC=0). Red spot marks the discrimination threshold point, in cases where R_{MAX} is a strong determinant of performance. In these cases, R_{MAX} deviates from its own pedestal level by an amount approximately equal to the noise level, representing a threshold level of discrimination performance, d'=1. In conditions 6 and 8, the red spot is absent because R_{LUSTRE} was the important cue (see Fig. S2). Note how in some tasks (7 and 11) sensitivity is quite poor (threshold ΔC is high) because potentially useful signals (B+, blue; L+, red) are largely vetoed by the first *MAX* operator.

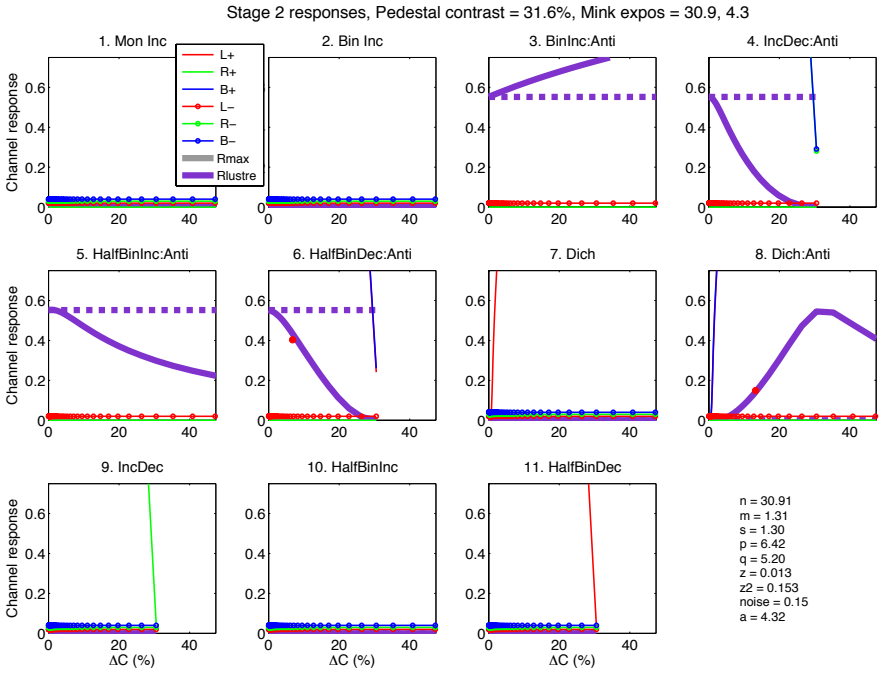

Figure S2. Diagnostic diagrams exactly like Fig S1, but plotting a low response range to illustrate R_{LUSTRE}. Dashed purple lines are the lustre response to pedestal-only. Six conditions (1,2,7,9-11) involved no negative polarity input, and so the negative-channel responses, along with R_{LUSTRE}, are zero. But in the five conditions (3,4,5,6,8) that involve antiphase contrasts, there were significant changes (increments or decrements) in the lustre response with increasing ΔC, and in two or three of these (conditions 6,8 and sometimes 4), lustre was the important cue for discrimination, according to the model.

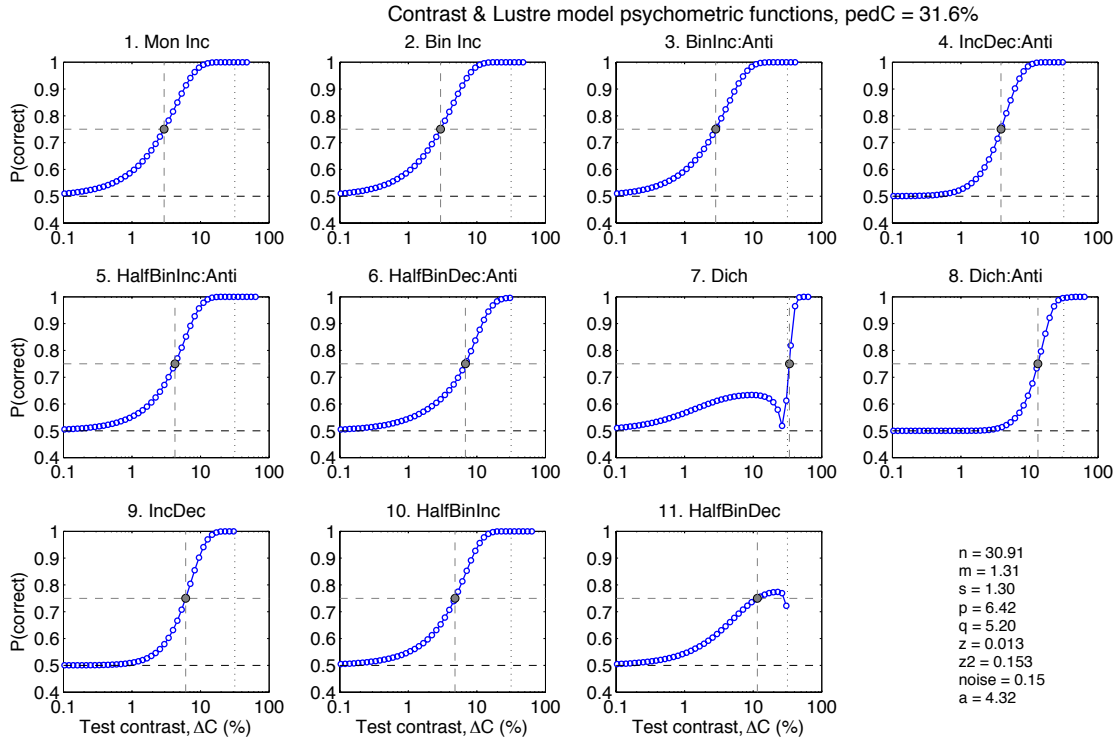

Figure S3. Model psychometric functions from the combined use of contrast and lustre responses shown in Figs. S1, S2. Pedestal contrast 32%. Grey spot marks the conventional discrimination threshold for ΔC (75% correct).

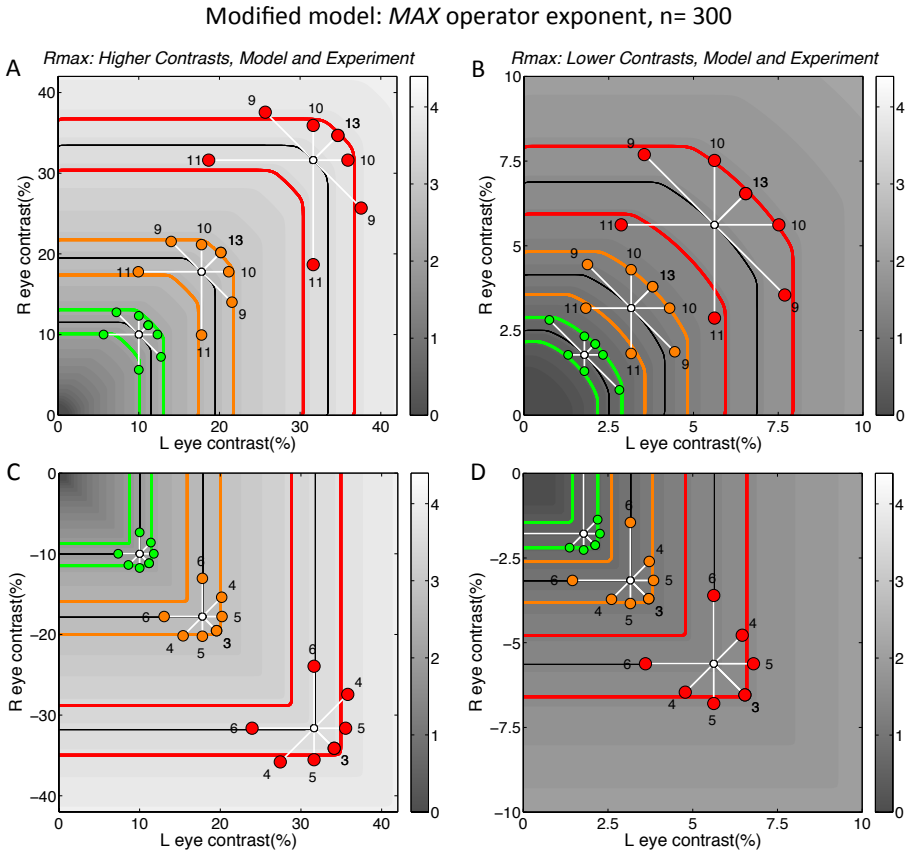

Figure S4. Why not use a perfect *max* operator ? This figure is similar to Fig. 6 in main text, but with one change to the model: parameter n = 300 instead of n = 31. This created an almost perfect max operator instead of the 'soft' or noisy max operator (Fig. A1). This gave the R_{MAX} response surface sharper corners. This generally had a minor influence, except for condition 11 at the two highest contrasts (panel A), where the surface was now so flat in the direction of change (white vector) that predicted performance saturated and could never reach threshold. In short, the 'soft max' operator (n=31) fits better, and could be interpreted as the operation of a 'hard' *max* along with noisy inputs.

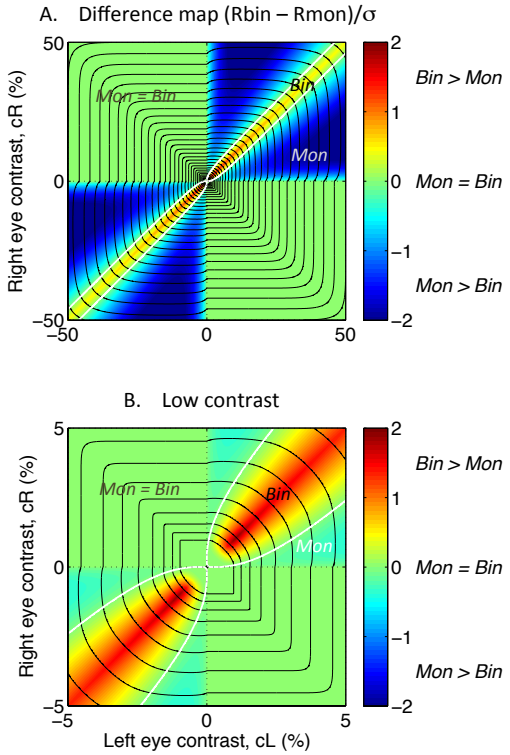

Figure S5. Difference maps show the distinct regions of binocular contrast space within which the model's monocular and binocular channels delivered the contrast cue, R_{MAX}. (A) bin-mon response differences mapped over a broad contrast range (0-50%), and (B) zoomed-in to low contrasts (0-5%). Yellow-red regions show where the binocular channel response was greater, and hence determined the value of R_{MAX}. Cyan-blue regions show where the monocular response was greater. White contour marks the boundary between them (i.e. bin-mon difference = 0). Note how at high contrasts (A) binocular responses were dominant only when left and right eye contrasts were nearly equal. At low contrasts (B) the influence of binocular responses expanded to cover a wide range of left-right contrast pairings. When polarities were opposite in the two eyes, mon and bin responses were equal everywhere in the 2nd and 4th quadrants (light green). **Defining the difference map:** *Rmon* was defined at each (c_L,c_R) point as the largest of the 4 monocular channel responses, i.e. $max(R_L^+, R_R^+, R_L^-, R_R^-)$, and similarly *Rbin* was defined as $max(R_B^+, R_B^-)$. The difference (colour-coded) is shown in units of the noise standard deviation, i.e. as $(Rbin-Rmon)/\sigma$. For reference, black contours are the iso-response contours of R_{MAX}, whose heights are spaced in steps of size σ.

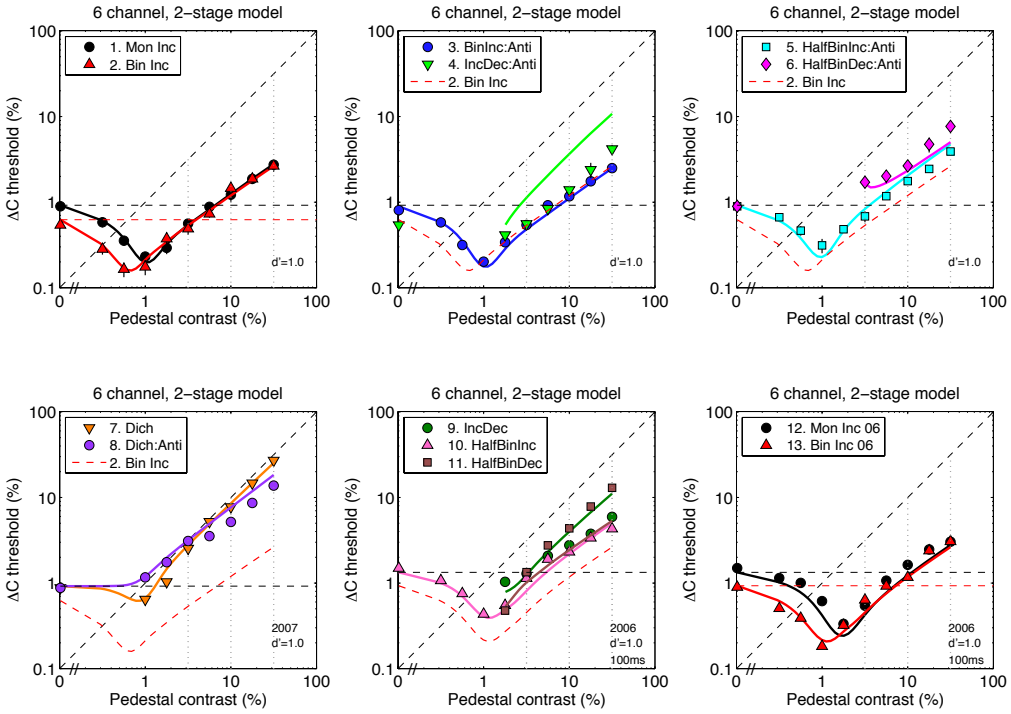

Figure S6. *Lustre* cue was necessary for a good model fit. Same as Figure 4 of the main text, except that the model was re-fitted with the response to *Lustre* removed (set to zero). Model performance now depended solely on the *Contrast* cue (R_{MAX}). Although overall R^2 was high (R^2=0.916, N=111; RMS error = 2.63 dB), this model showed a poor fit to the data in conditions 4, 9, 11, and to some extent in 5, 6, 8. An F-test comparing the fits of the two nested models (nine free parameters with *Lustre vs* eight without *Lustre*) was hugely significant [F(1,102) = 422.9, P<0.00001], meaning that *Lustre* significantly improved the fit. Best-fitting parameters with no *Lustre* were: n = 6.096; m = 1.396; s = 1.642; p = 5.275; q = 4.340; z = 0.021; σ = 0.144; a = (irrelevant; no effect); z_2 = 0.167.

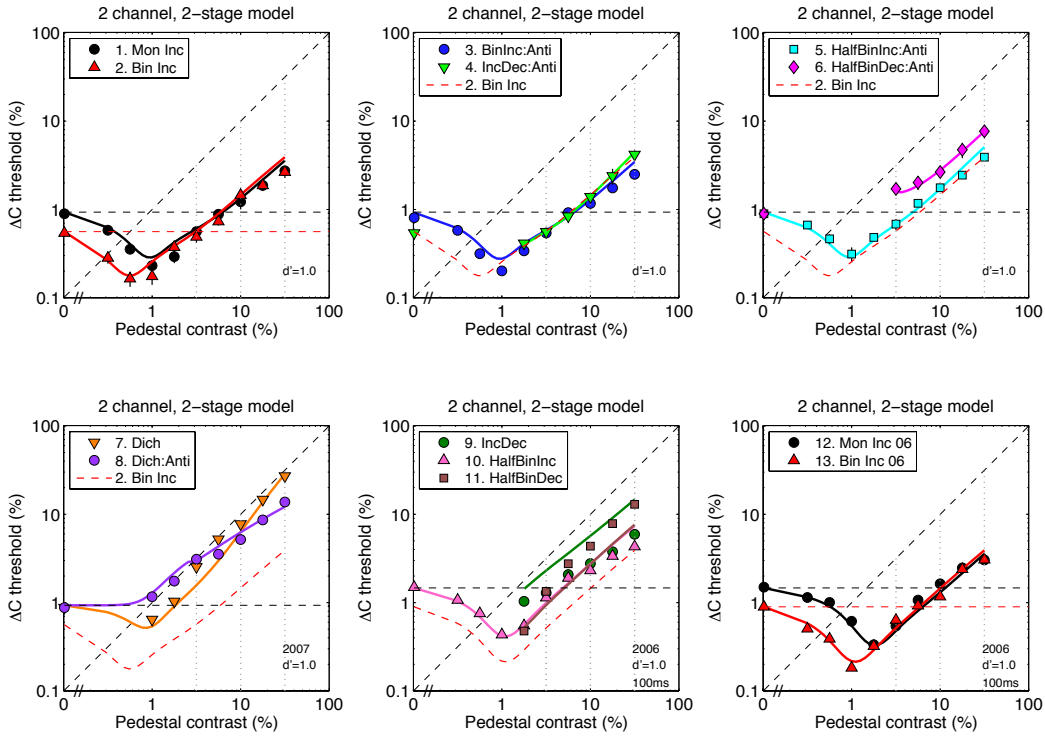

Figure S7. Monocular channels were needed for a good fit to the whole dataset. Like Fig. 4 of the main text, except that the model was re-fitted with monocular channel responses set to zero at stage 2. Model performance depended on the *Contrast* and *Lustre* cues delivered by the binocular channels. Overall R^2 was high (R^2=0.942, N=111; RMS error = 2.19 dB), and this model showed a good fit to the data in all conditions except 9 and 11 where the fit was poor. No nested F-test comparison with the full model was possible in this case, because both models had the same number of free parameters, and differed only in the presence or absence of the monocular channels. Best-fitting parameters without monocular channels were: n=32.25; m=1.29; s=4.25; p=4.50; q=3.58; z=0.0023; σ=0.0984; a=4.09; z_2=0.027.

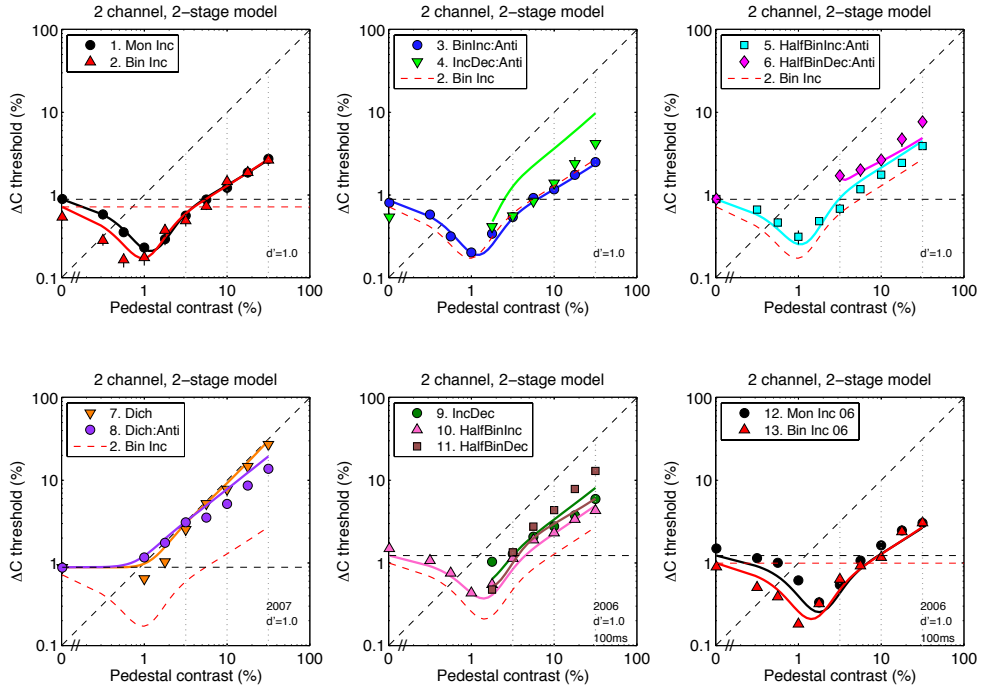

Figure S8. To confirm the effects shown in Figures S6, S7 the model was re-fitted with both changes together: no *Lustre* and no monocular channels. Model performance depended solely on the *Contrast* cue (R_{MAX}) delivered by binocular channels. Overall R^2 was high (R^2=0.917, N=111; RMS error = 2.61 dB), but this model again showed a poor fit to the data in conditions 4, 9, 11, and to some extent in 5, 6, 8, rather similar to the effect of removing *Lustre* alone (Fig. S6). An F-test comparing the fits of the two nested models (this one *vs* the full model of Fig. 4) was hugely significant [F(1,102) = 415.0, P<0.00001]. Best-fitting parameters were: n=6.301; m=3.186; s=28.56; p=1.221; q=1.055; z=0.0696; σ=0.119; a=(irrelevant; no effect); z_2=0.253.

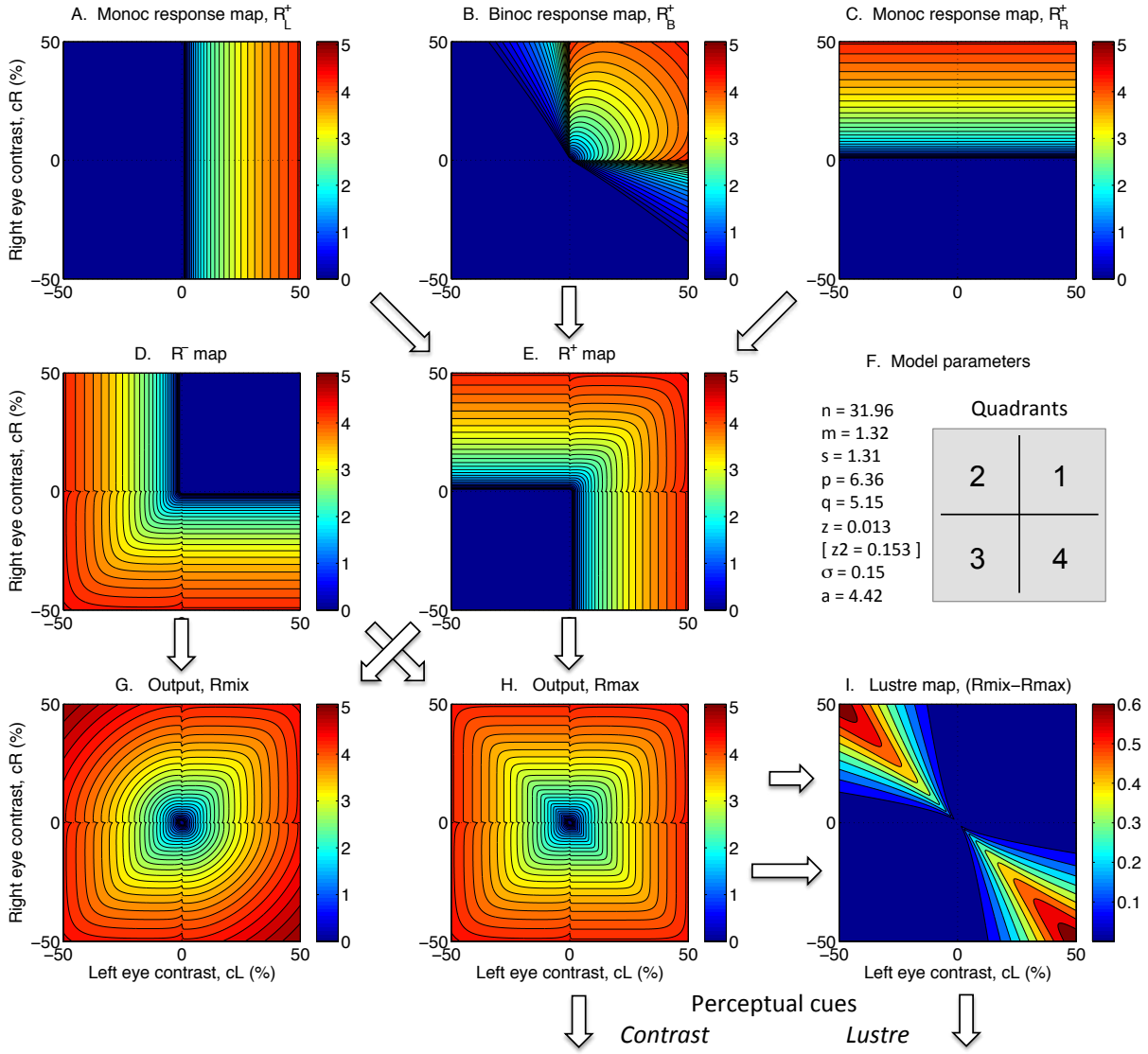

Figure S9. Model response surfaces, like Fig. 5 of the main text - except that the model was re-fitted with binocular channels that responded to the difference between opposite-polarity inputs (Eqn. 7a). In panel B, note the cancellation of binocular-channel responses to inputs that have opposite-polarity (quadrants 2 and 4). However, because of the *max* operator, the monocular responses win the day, and this cancellation does not carry through to later response stages (D,E,G,H), and has almost no effect on fitted parameters (F) or predicted thresholds. We therefore cannot determine from the data whether such cancellation occurs or not; see Sec 5.5 of main text.