

# Binocular integration of chromatic and luminance signals

Daniel H. Baker, Kirralise J. Hansford, Federico G. Segala, Annie Y. Morsi,  
Rowan J. Huxley, Joel T. Martin, Maya Rockman & Alex R. Wade

Department of Psychology, University of York

## 1 Abstract

## 2 Introduction

The process by which the brain combines independent inputs is of fundamental importance for understanding sensory perception. Binocular vision is a useful test-case for determining the general principles involved in neural signal combination, as our brains typically combine the inputs from the left and right eyes to provide binocular single vision. In recent years our understanding has been facilitated by the development of binocular gain control models that provide a framework to interpret empirical data from multiple experimental paradigms and techniques, including psychophysics (Meese et al., 2006), EEG (Baker & Wade, 2017), fMRI (Moradi & Heeger, 2009) and pupillometry (Segala et al., 2023). However, the majority of this work has used achromatic (black and white) stimuli; we know comparatively little about how chromatic signals are combined binocularly, or about how signals in different ocular and chromatic channels interact. In this study we use psychophysical detection and discrimination paradigms to explore binocular interactions within and between the chromatic and achromatic pathways.

A useful framework for understanding binocular signal processing is the two-stage gain control model of binocular combination introduced by Meese et al. (2006). This model features interocular suppression between monocular channels, followed by binocular summation. The model accounts well for the pattern of contrast discrimination (‘dipper’) functions for four distinct ocular configurations (see also Georgeson et al., 2016), illustrated in Figure 1. In the monocular condition, participants must discriminate between stimuli of two contrasts (a ‘pedestal’, and a ‘pedestal plus target’) that are both presented to one eye, whilst the other eye views mean luminance. Threshold is defined as the minimum target contrast required to make this judgement with 75% accuracy. The binocular condition is the same, except that the same stimuli are shown to both eyes. In the half-binocular condition the pedestal is shown to both eyes, but the target increment shown only to one eye. Finally, the dichoptic condition involves presenting the pedestal to one eye, and the target increment to the other eye.

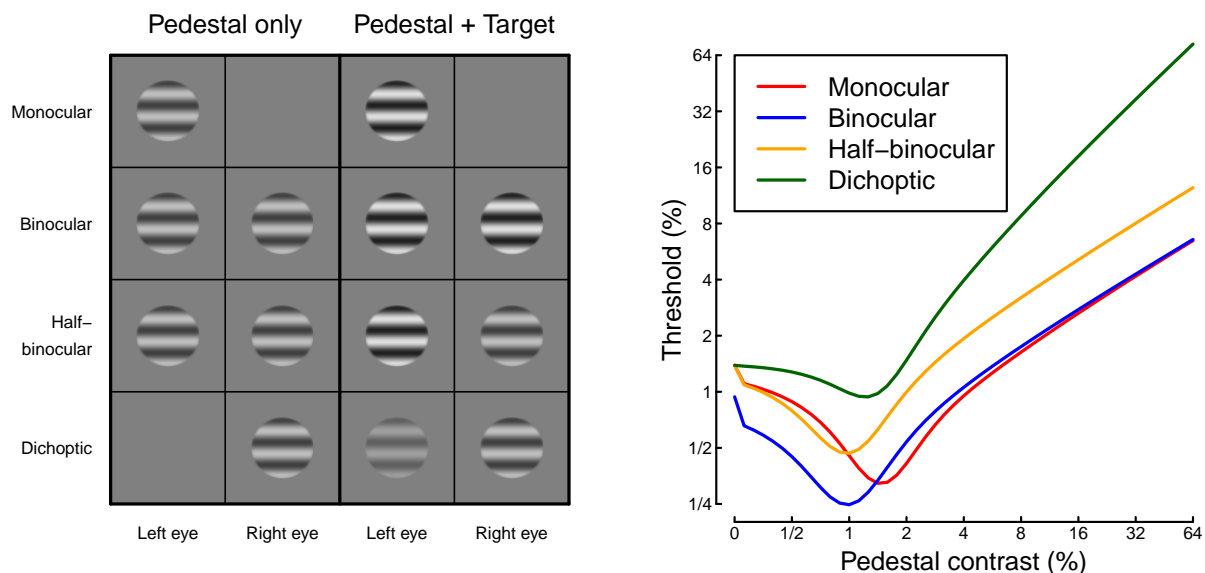


Figure 1: Illustration of stimulus conditions (left) and example dipper functions (right).

The detailed pattern of thresholds across these four conditions is complex, and for achromatic stimuli has several distinctive features that have been replicated in multiple studies. At low pedestal contrasts, the binocular condition yields lower thresholds than the monocular condition; this result is attributed to physiological binocular summation by neurons responsive to signals from both eyes (Baker et al., 2018; Campbell & Green, 1965). However at high pedestal contrasts the ‘handle’ regions of the dipper functions for these conditions converge: a consequence of interocular suppression compensating for the increased excitation during binocular stimulation (Legge, 1984; Maehara & Goryo, 2005). The half-binocular condition avoids confounding the number of eyes seeing the target with the number of eyes seeing the pedestal (Meese et al., 2006). The pedestal is always binocular in this condition, whereas the target increment is monocular, and thresholds are consistently higher than in the binocular condition across the full range of pedestal contrasts. This demonstrates that binocular summation occurs across the full contrast range, when the pedestal ocularity is appropriately controlled. Finally, the dichoptic condition produces extremely strong masking of the target, such that when the pedestal is visible, the target must equal or exceed its contrast in order to be detectable (Baker & Meese, 2007; Legge, 1979; Maehara & Goryo, 2005). The characteristic pattern of dipper functions is well-described by the gain control model of Meese et al. (2006) for achromatic stimuli.

At the output of the human retina, cone responses are split into three distinct pathways. The sum of long- and medium-wavelength cone outputs (L+M) transmits luminance information, and is likely responsible for the binocular combination effects previously studied using achromatic stimuli (see above). The difference of long- and medium-wavelength cone outputs (L-M) is responsive to chromatic stimuli modulating along a red/green axis in colour space. Finally the short wavelength cone outputs (S-(L+M)) code chromatic stimuli modulating along a blue/yellow axis. There has not yet been a detailed investigation of binocular contrast discrimination in either of these chromatic pathways, however there is reason to believe they may differ from the achromatic pathway. At detection threshold, binocular summation is greater for chromatic versus achromatic stimuli (Simmons, 2005), implying a more linear initial stage of processing. For cross-orientation masking, there are differences in the magnitude of masking between chromatic (red/green) and achromatic stimuli (Kim et al., 2013; Medina & Mullen, 2009), as well as differences in their temporal dynamics (Kim & Mullen, 2015). There are also interactions between chromatic and achromatic pathways both within (Chen et al., 2000) and between (Kingdom & Libenson, 2015; Mullen et al., 2014) the eyes, yet these have not been fully explored for arrangements where the target and mask have the same orientation. Finally, the

neurophysiological underpinnings of colour vision are distinct from those of the achromatic system. In primary visual cortex (V1), chromatic signals are processed in ‘blob’ regions that are revealed by cytochrome oxidase staining (Horton & Hubel, 1981). The blob regions appear to be largely monocular (Livingstone & Hubel, 1984), suggesting that binocular combination for chromatic stimuli might occur later than for achromatic stimuli, and perhaps be subject to different constraints.

A further property of ‘blob’ regions is that they are less strongly orientation-tuned than other regions of V1 (Horton & Hubel, 1981; Livingstone & Hubel, 1984), and biased towards low spatial frequencies (Edwards et al., 1995; Tootell et al., 1988). Our recent work has investigated binocular combination for flickering discs of luminance, which are DC-balanced across time (Segala et al., 2023). Steady-state EEG responses from early visual cortex and psychophysical contrast matching data were both consistent with weak interocular suppression when using this stimulus arrangement. This is very different from the ‘ocularity invariance’ that is well-established when using DC-balanced periodic stimuli such as sine-wave gratings, and implies strong interocular suppression (Baker & Wade, 2017; Meese et al., 2006; Moradi & Heeger, 2009). There is additional psychophysical evidence from matching studies that binocular combination can be close to linear for luminance increments (Anstis & Ho, 1998; Levelt, 1965), particularly against a dark background (Baker et al., 2012). Both structural and functional data therefore imply that binocular combination may differ between spatial and temporal contrast.

The main aim of the present study is to characterise binocular signal combination for chromatic stimuli, and for temporal modulations of luminance. We also aimed to investigate interocular suppression between chromatic and achromatic pathways. We therefore preregistered a series of psychophysical experiments (see: <https://osf.io/3vdga/>). In Experiment 1 we replicate the four key pedestal masking conditions of Meese et al. (2006) described above for achromatic grating stimuli, and extend this to both red/green and blue/yellow isoluminant chromatic stimuli. In Experiment 2 we explore dichoptic masking within and between these stimuli. Experiment 3 repeats the achromatic condition from the first experiment, but using a temporally modulated disc rather than sine-wave gratings. We take a Bayesian approach to data analysis and modelling; by fitting a hierarchical version of the two-stage gain control model (Meese et al., 2006) we compare posterior parameter distributions to understand how model parameters such as the weight of interocular suppression vary across visual pathways.

## 3 Materials & Methods

### 3.1 Participants

All experiments were completed by the first author (DHB) and two additional participants, who differed for each experiment. Participants had no known abnormalities of binocular vision. Written informed consent was obtained before data collection began, and all procedures were approved by the ethics committee of the Department of Psychology at the University of York (ID number 2202).

### 3.2 Apparatus & stimuli

In Experiments 1 and 2, the stimuli were horizontal sinusoidal gratings with a spatial frequency of 1c/deg (see examples in Figure 2). The gratings were windowed by a raised cosine envelope with a diameter of 3 degrees. Spatial phase, relative to a central fixation cross, was randomised on each trial across the four cardinal phases. In the achromatic conditions, the sine-wave modulated all three monitor colour channels equally (red, green and blue). In the L-M condition, we generated isoluminant stimuli for each participant (see Procedures) designed to maximise contrast between L and M cones, whilst keeping S cone activity constant. In the S-(L+M) condition, the isoluminant stimuli maximised S-cone contrast. Stimuli were converted from cone space to monitor RGB coordinates using the monitor spectral readings and the Stockman-Sharpe 2 degree cone fundamentals (Stockman & Sharpe, 2000). The stimuli in Experiment 3 were temporal modulations of luminance applied to a disc made using the same raised cosine envelope as described above, but with no further spatial modulation. The stimuli counterphase flickered sinusoidally at 4Hz (see examples in the lower portion of Figure 2). In all experiments, we displayed a binocular fusion lock, consisting of three

concentric rings of small square elements with random colour. A black central fixation cross was also displayed throughout.

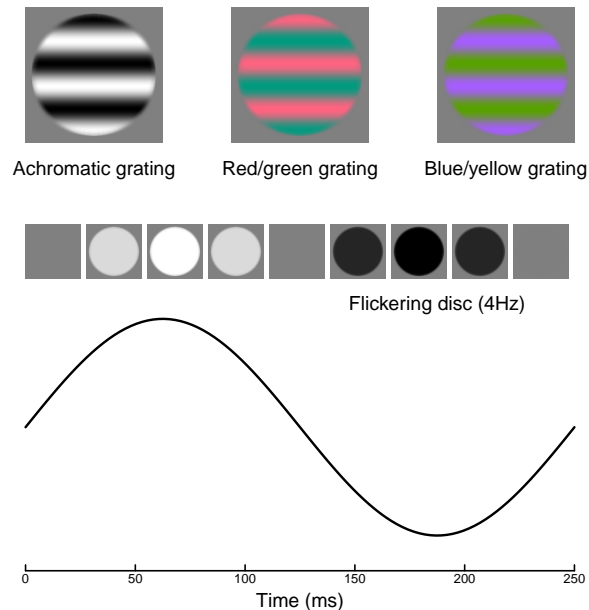


Figure 2: Example stimuli. Upper row shows grating stimuli used in Experiments 1 and 2. Lower row shows one cycle of sinusoidal flicker applied to a uniform disc. Note that the rendering of all stimuli will depend on the device used to display or print this image, and so the chromatic stimuli are unlikely to appear isoluminant, and there may be additional luminance nonlinearities that were not present in the stimuli displayed during the experiments.

All stimuli were presented on an Iiyama VisionMaster Pro 510 CRT monitor, with a refresh rate of 100Hz, and a resolution of 1024 x 768 pixels. The display was driven by a ViSaGe MkII stimulus generator (Cambridge Research Systems Ltd., Kent, UK) running in 42-bit colour mode (14 bits per colour channel). We presented stimuli to the left and right eyes independently using a four-mirror stereoscope with front-silvered mirrors. The display was luminance calibrated using a ColourCal photometer (Cambridge Research Systems), and gamma corrected by fitting a four-parameter gamma function to the output of each CRT gun. The maximum luminance was 87 cd/m<sup>2</sup>. We also measured the spectral output of each phosphor using a Jaz spectroradiometer (Ocean Insight, Florida), and used these measurements to convert between LMS (cone) space and the monitor RGB coordinates.

For convenience, we express stimulus contrast as a percentage of the maximum possible contrast that could be displayed on our system. For achromatic stimuli, the maximum contrast is 1, so this is equivalent to the standard Michelson contrast expressed as a percentage. For the isoluminant chromatic stimuli, the maximum displayable L-M (red/green) cone contrast was 0.1, and the maximum displayable S-(L+M) (blue/yellow) cone contrast was 0.88. An L-M threshold of 50% therefore corresponds to a cone contrast of  $0.5 \times 0.1 = 0.05$ , and an S-(L+M) threshold of 50% corresponds to a cone contrast of  $0.5 \times 0.88 = 0.44$ . This means that the threshold values reported throughout can be converted to cone contrast by a straightforward multiplicative transform.

### 3.3 Procedure

All experiments took place in a darkened room. Participants placed their heads in a chin rest mounted on a height-adjustable table, to which the stereoscope was also attached. The total optical viewing distance

(including the light path through the mirrors) was 104cm, at which distance 1 degree of visual angle encompassed 48 pixels on the monitor.

Before beginning primary data collection, each participant in Experiments 1 and 2 completed an isoluminance adjustment task. Grating stimuli were presented that counterphase flickered at 5Hz, defined about either the L-M or S-(L+M) plane in cone space. Participants used a trackball to dynamically adjust the colour angle of the stimulus to minimise the percept of flicker. Each participant completed ten such trials for each colour plane, and the average angle across repetition was taken as the isoluminant point, and used to generate stimuli for the main experiment for that participant. Settings were very similar across participants for the S-(L+M) direction, and somewhat more heterogeneous for the L-M direction (see Figure 3).

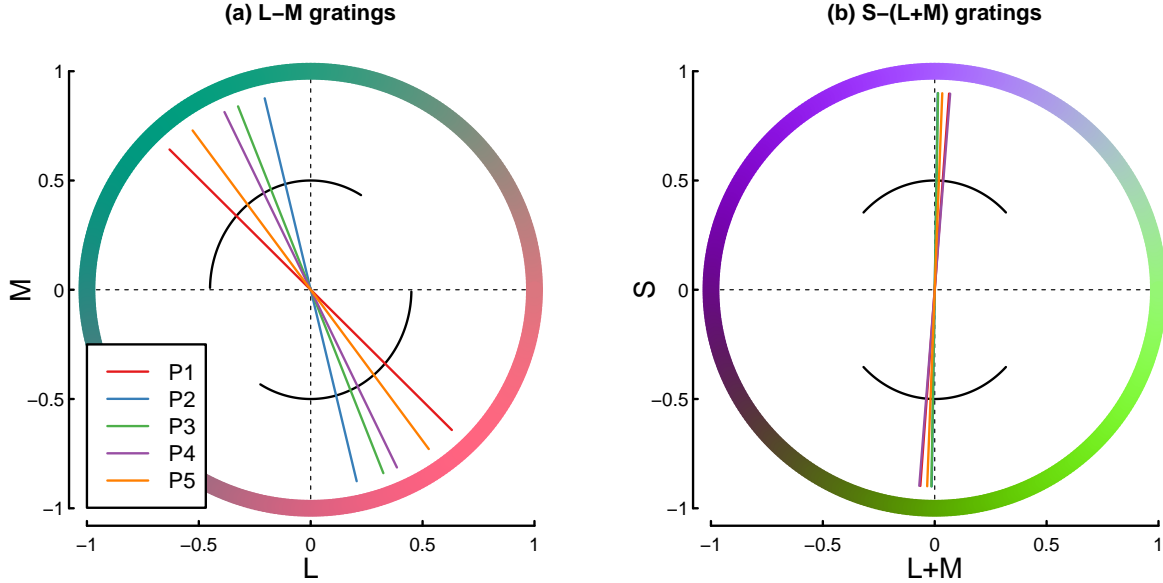


Figure 3: Isoluminance settings from all participants in Experiments 1 and 2. Panel (a) shows red/green and panel (b) shows blue/yellow settings that were subsequently used to generate stimuli in the main experiments. Within each panel, solid lines show the mean settings for each participant, and black curves show the range of possible stimuli displayed during the adjustment task.

In Experiment 1, participants completed a two-interval-forced-choice (2IFC) contrast discrimination task. Stimuli were presented for 200ms, with an interstimulus interval of 400ms. Each interval was indicated by an auditory beep, and participants made their responses using a two-button trackball. Correct responses were followed by a high pitched tone, and incorrect responses by a low pitched tone. Each block of the experiment tested a single pedestal contrast level, and lasted around 12 minutes. On each trial the target contrast level was determined by a 3-down-1-up staircase procedure. There were 8 interleaved staircases in total; four stimulus arrangements (see Figure 1) combined factorially with two target eye assignments. Each pedestal contrast was repeated 3 times by each participant, and the block order was randomised. The experiment lasted around 4 hours per participant for each chromatic condition, and took place over the course of several weeks. In total, the experiment consisted of 67978 trials (pooled across participants).

In Experiment 2, participants completed a 2IFC dichoptic masking task. The stimuli and trial protocol were the same as for Experiment 1, except that the target contrast was chosen from a set of 10 possible values, determined in advance based on the data of Experiment 1. There were 12 possible conditions: baseline detection thresholds for achromatic, red/green and blue/yellow stimuli, and the nine possible factorial pairings obtained by assigning these conditions to be target and dichoptic mask stimuli. Mask contrasts were chosen

to be approximately 16 times their (monocular) detection threshold, based on the data from Experiment 1. As a percentage of the maximum displayable contrast, these were 16% for achromatic stimuli, and 64% for both of the chromatic stimuli. Each block of the experiment tested a single condition, and consisted of 200 trials. A high contrast example of the target stimulus was displayed at the foot of the screen throughout, so that there was no ambiguity about the target identity on a given block. Participants completed 10 repetitions of each condition (120 blocks of ~6 minutes each), lasting around 12 hours, for a total of 72000 trials (pooled across participants).

In Experiment 3, the achromatic conditions from Experiment 1 were repeated using a flickering disc stimulus. The stimulus counterphase flickered at 4Hz, and was presented for 500ms (i.e. 2 full cycles of the temporal modulation). All other procedures were the same as for Experiment 1, and the experiment comprised a total of 24610 trials (pooled across participants).

### 3.4 Data analysis and computational modelling

Psychometric functions from each experiment were fit using *psignifit* 4 to estimate threshold and slope parameters via a Bayesian numerical integration method (Schütt et al., 2016). A cumulative Gaussian was used as the underlying function, and we converted the slope estimates ( $\sigma$  parameters from the fitted Gaussians) to equivalent Weibull  $\beta$  values using the approximation  $\beta = 10.3/\sigma$ .

The two stage model of Meese et al. (2006) was fit to the threshold data from Experiments 1 and 3 using a simplex algorithm to minimise the error between the model and data. The model is defined by a series of equations:

$$Stage1_L = \frac{C_L^m}{S + C_L + \omega C_R}, \quad (1)$$

$$Stage1_R = \frac{C_R^m}{S + C_R + \omega C_L}, \quad (2)$$

$$binsum = Stage1_L + Stage1_R, \quad (3)$$

$$Stage2 = \frac{binsum^p}{Z + binsum^q}, \quad (4)$$

where  $C_L$  and  $C_R$  are the contrasts displayed to the left and right eyes, and  $m$ ,  $S$ ,  $\omega$ ,  $p$ ,  $q$  and  $Z$  are free parameters in the model. A further free parameter,  $k$ , represents additive internal noise, and is used to convert the model outputs to either d-prime or threshold values (note that in the original model specification (Meese et al., 2006) this parameter was called  $\sigma$ , but we use the  $k$  symbol here to avoid confusion with the standard deviation of the cumulative Gaussian used when fitting the psychometric functions to derive thresholds). Thresholds are defined by iteratively adjusting the target contrast until the following equality is satisfied:

$$Stage2_{target+pedestal} - Stage2_{pedestal} = k, \quad (5)$$

and d-prime for a single target level is defined as:

$$d' = \frac{Stage2_{target+pedestal} - Stage2_{pedestal}}{k}. \quad (6)$$

We ran the simplex algorithm from 100 random starting vectors for each data set, and chose the solution for each data set that gave the smallest RMS error.

We also implemented a Bayesian hierarchical version of the model using the Stan language (Carpenter et al., 2017). This used a binomial generator function to model the proportion correct data at each target level, and was fit simultaneously to all participants for a given experiment, but separately for each chromatic condition of Experiment 1, and the flickering disc data from Experiment 3 (i.e. four fits in total, as for the simplex fitting). Prior distributions were Gaussian, with means determined from published values (see first row of Table 1). This modelling primarily focuses on examining posterior parameter distributions, rather than a model comparison approach. We generated over 1 million posterior samples for the model, and retained 10% of them for plotting.

Finally, we adapted the two stage model to include parallel pathways to process achromatic, red/green and blue/yellow stimuli, that mutually suppress each other. We added additional suppressive terms at the first (monocular) stage of the model, for example:

$$ACStage1_L = \frac{AC_L^m}{S + AC_L + \omega_A AC_R + \omega_R RG_R + \omega_B BY_R}, \quad (7)$$

where  $AC$  represents the achromatic contrast,  $RG$  represents the red/green contrast,  $BY$  represents the blue/yellow contrast, and  $\omega_A$ ,  $\omega_R$  and  $\omega_B$  are the accompanying weights of interocular suppression. There is an equivalent expression for the right eye, and for each of the two isoluminant chromatic pathways. To simplify the model, and avoid free parameters that are poorly constrained by the data, we fixed several parameters at the average values from the fits from Experiment 1 (Table 1), such that  $p = 7.82$ ,  $q = 6.4$ ,  $m = 1.21$ ,  $S = 1.03$  and  $k = 0.23$ . We additionally fixed the within-mechanism weight of interocular suppression ( $\omega$ ) at 1. This left 9 free parameters: a  $Z$  parameter for each mechanism, and six cross-mechanism weights of interocular suppression. These parameters were again estimated within a Bayesian hierarchical framework, using the trial-wise data from Experiment 2. Note that the model as specified does not currently include monocular suppression between different pathways, as we did not collect any data for these conditions. Previous work (Chen et al., 2000) has measured such interactions, and they could in principle be incorporated into the denominator of either stage 1 or stage 2 in the model.

### 3.5 Open science practices

All experimental code, raw data and analysis scripts are available at: <https://osf.io/3vdga/>. The linked GitHub repository also contains a fully reproducible version of the manuscript. Note that we deviated slightly from the planned preregistration, in that we did not collect data for chromatic flickering discs, or for the cross-pathway dichoptic experiment using disc stimuli. This is because the grating data from Experiments 1 and 2, and the achromatic disc data from Experiment 3, were sufficient to address the questions we had hoped to answer from these experiments.

## 4 Results

### 4.1 Experiment 1

Dipper functions from Experiment 1 are displayed in the upper row of Figure 4. Panel (a) shows the achromatic results, which replicate the key features from previous work. At detection threshold, binocular summation was a factor of 1.67 (4.47dB), within the range ( $\sqrt{2}$  to 2) consistent with previous reports (Baker et al., 2018). Pedestal masking functions followed the typical ‘dipper’ shape in all conditions, with a region of facilitation at low pedestal contrasts, and masking at higher contrasts. The monocular and binocular dipper handles converge at high contrasts, whereas the half-binocular thresholds remain above the binocular thresholds across the full range of pedestal contrasts. The dichoptic condition produced very high thresholds, with the rising portion of the dipper having a slope around 1 (regression slope of 1.06 in log (dB) units, calculated across the highest 4 pedestal contrasts).

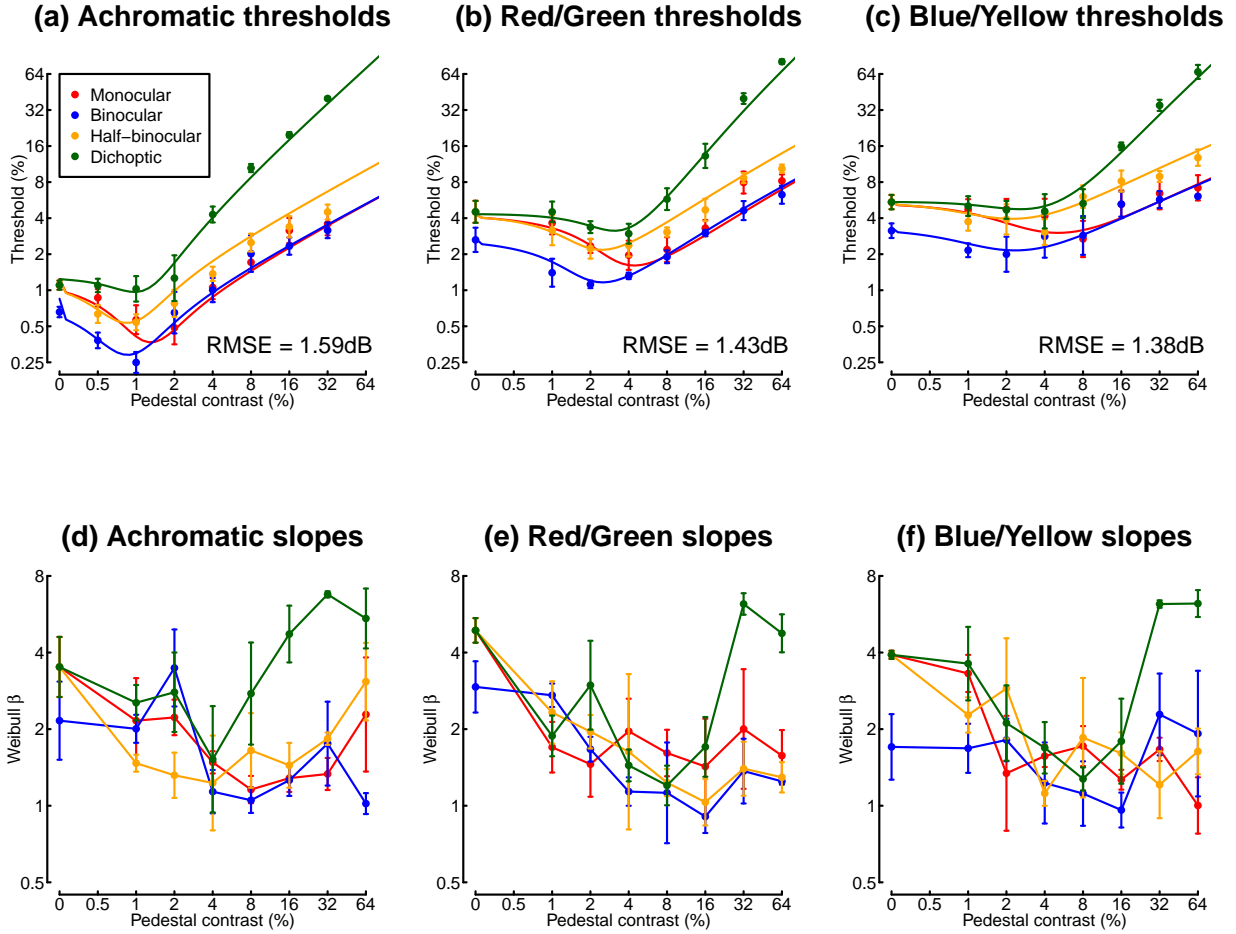


Figure 4: Dipper functions and psychometric slopes from Experiment 1, averaged across three participants. Panels (a-c) show threshold data, and panels (d-f) represent the slope of the psychometric function expressed in Weibull  $\beta$  units. Error bars give  $\pm 1$ SE across participants. Note that contrast values are expressed as a percentage of the maximum displayable contrast (see Procedures for details). Curves in panels (a-c) show the best fitting models, optimized using a simplex algorithm (described in section 4.4), and RMSE values give the root mean square errors of the fits.

A similar pattern of results was observed for both the red/green and blue/yellow isoluminant stimuli (see Figure 4b,c). Summation at threshold was a factor of 1.71 (4.66dB) for the red/green targets, and a factor of 1.73 (4.77dB) for the blue/yellow targets, and so was marginally higher than for achromatic stimuli. The general character of the dipper functions was largely consistent with the achromatic results, though we observed shallower facilitation and weaker masking, especially for the blue/yellow stimuli. For example, the strongest facilitation in the binocular condition for achromatic stimuli was a factor of 2.63, whereas it reduced to a factor of 2.35 for red/green stimuli and 1.57 for blue/yellow stimuli. The slope of the binocular dipper handle was 0.52 for achromatic stimuli, 0.58 for red/green stimuli, and 0.34 for blue/yellow stimuli. Dichoptic masking remained as strong for the chromatic conditions as for the achromatic stimuli (regression slopes of 1.3 for red/green and 1.21 for blue/yellow). The pattern of results for individual participants was consistent with the group averages, as shown in Figure A1.

Following Meese et al. (2006), we also inspected the slope of the psychometric function for each condition



(see Figure 4d-f). At detection threshold, slopes were relatively steep for all stimuli, with  $\beta$  values around 4. As pedestal contrasts increased, slopes linearized and reduced to around  $\beta = 1.3$  (Foley & Legge, 1981; Meese et al., 2006), and remained shallow at high pedestal contrasts. The exception to this was the dichoptic condition, where slopes became extremely steep at high dichoptic mask contrasts, consistent with previous observations (Baker et al., 2013; Meese et al., 2006). This was clear for all three data sets, with slope values in the range  $4 < \beta < 8$ . However, we observe that slope estimates are more variable than threshold estimates (note the large error bars), particularly when using adaptive staircases, which deploy the majority of trials close to threshold. Our second experiment therefore investigated the slope of the psychometric function in more detail for dichoptic masking using the method of constant stimuli.

## 4.2 Experiment 2

In Experiment 2 we focussed on the dichoptic condition at a single mask contrast, and measured full psychometric functions using the method of constant stimuli for all factorial pairings of target and mask chromaticity. The pooled results across three participants are shown in Figure 5a-c, and results for individual participants are available in Figure A2. All conditions produced monotonically increasing psychometric functions (panels a-c), but the extent of masking was highly dependent on the relationship between the target and mask chromaticity. Figure 5d-f shows a two-dimensional representation of individual threshold and slope estimates (points), as well as the posterior density estimates for fits to the pooled data (ellipses). The results are consistent between participants, and at the group level, and show that the presence of a mask has a strong effect on thresholds.

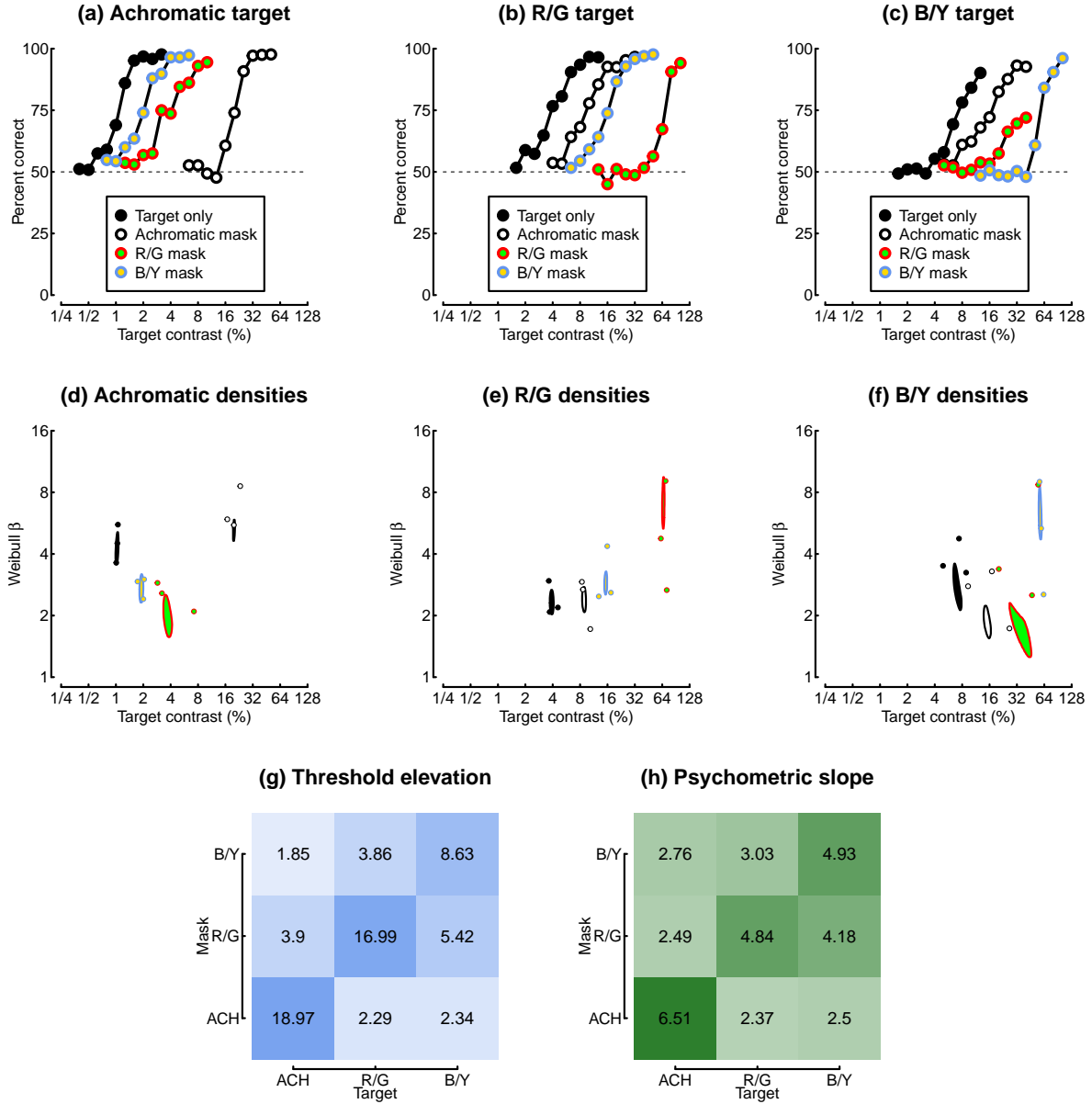


Figure 5: Summary of data from Experiment 2. Panels (a-c) show psychometric functions for each condition, pooled across participants (600 trials per target contrast level). Panels (d-f) show threshold and slope estimates for individual participants (points) and the boundary of the posterior density estimates for fits to the pooled data (ellipses). Panel (g) shows the average threshold elevation factor for each combination of target and mask stimulus. Panel (h) shows the geometric mean psychometric slope value for each masking condition in Weibull  $\beta$  units.

Threshold elevation was greatest when the target and mask had the same chromaticity - notice that the psychometric function is shifted furthest to the right for the achromatic target with an achromatic mask (white and black circles in Figure 5a), for the red/green target with a red/green mask (red and green circles in Figure 5b), and for the blue/yellow target with a blue/yellow mask (blue and yellow circles in Figure 5c). Masking was weakest between achromatic masks/targets and chromatic masks/targets. Finally there was an intermediate level of masking between red/green and blue/yellow stimuli. This is summarised in Figure 5g,

which represents threshold elevation for each combination of target and mask chromaticity. Note that the positive diagonal exhibits the highest values, and represents threshold elevation between targets and masks of the same chromaticity.

We also calculated the slope of the psychometric function for each condition, in equivalent Weibull  $\beta$  units. In the absence of a mask, the average slope was  $\beta = 3.42$ , which is typical for contrast detection tasks (Wallis et al., 2013). Slopes became substantially steeper when the dichoptic mask matched the target in chromaticity (average  $\beta = 5.38$ ). These ‘super-steep’ psychometric functions for dichoptic pedestal masking have been reported previously (Baker et al., 2013; Meese et al., 2006), and are observed for the first time here using chromatic stimuli (see diagonal values in Figure 5h, and also Figure 4d-f). However we did not see such markedly steep functions for any of the cross-chromaticity masking conditions (average  $\beta = 2.83$  for the off-diagonal values).

### 4.3 Experiment 3

In our final experiment, we again measured dipper functions, but this time for a temporally modulating luminance disc. This was motivated by our recent work (Segala et al., 2023) that appeared to show increased binocular facilitation and reduced interocular suppression for flickering disc stimuli (relative to gratings), measured using EEG and a psychophysical matching paradigm. The pattern of dipper functions for a 4Hz flickering disc (see Figure 6a) was very similar to that observed for achromatic gratings (see Figure 4a), and the binocular summation ratio at threshold was also similar (a factor of 1.49 for discs, vs 1.67 for gratings). Threshold data for individual participants are shown in Figure A3. We also found a similar pattern of psychometric slope values (Figure 6b) as we had for gratings, though we note that the dichoptic condition did not produce the ‘super-steep’ psychometric functions we had observed in Experiments 1 & 2. Nevertheless, the dichoptic slopes are somewhat above those of the other pedestal arrangements at high contrasts.

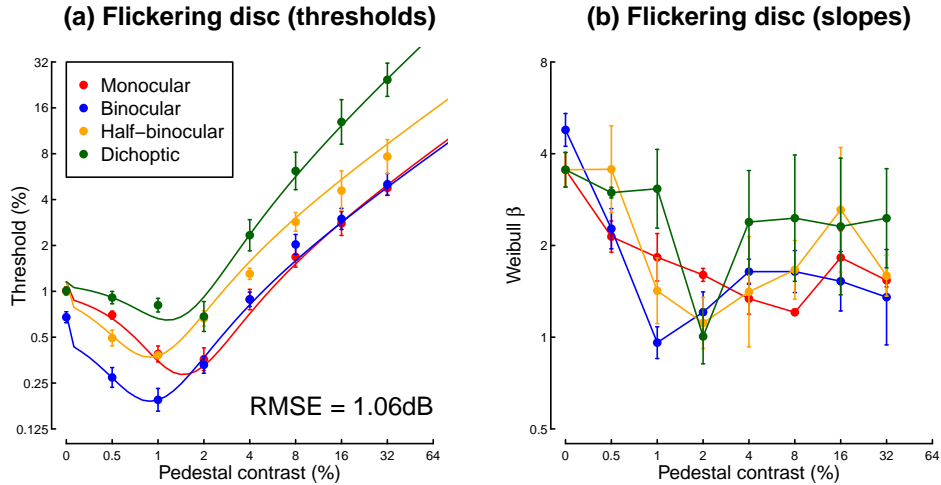


Figure 6: Thresholds for the flickering disc experiment. Plotting conventions mirror those in Figure 4.

### 4.4 Computational modelling

For consistency with previous work, we initially performed least-squares fits of the two-stage model (Meese et al., 2006) for each of the chromaticity experiments, and the flickering disc experiment (7 free parameters per fit). The best model fits are shown by the curves in Figures 4 & 6, and provide an excellent description of the data, with RMS errors between 1.06 and 1.59dB. Best fitting parameters are shown in the ‘Simplex fits’ section of Table 1. We note that in previous work, the weight of interocular suppression ( $\omega$  in the model) is implicitly fixed at 1. Here we allowed it to vary, but it still received a value of exactly 1 in all of our

Table 1: Summary of fitted model parameters. The top row gives the best fitting parameters from the study of Meese et al. (2006). The second section shows the best fitting parameters from simplex fits to the averaged thresholds for each experiment. The final rows show the posterior parameter estimates from the Bayesian model, fitted to each data set from Experiments 1 and 3.

<i>Model fit</i>	<i>p</i>	<i>q</i>	<i>m</i>	<i>S</i>	<i>Z</i>	$\omega$	<i>k</i>	<i>RMSE</i>
Meese et al. (2006)	7.99	6.59	1.28	0.99	0.08	1.00	0.19	
<b>Simplex fits</b>								
Achromatic gratings	6.51	5.10	1.31	0.83	0.16	1.00	0.21	1.59 dB
Red/green gratings	19.42	1.10	1.03	0.37	1.49	1.00	0.29	1.43 dB
Blue/yellow gratings	9.47	6.21	1.14	1.89	0.26	1.00	0.34	1.38 dB
Flickering discs	6.13	4.93	1.24	0.84	0.27	0.89	0.13	1.06 dB
<b>Bayesian model</b>								
Achromatic gratings	7.78	6.37	1.29	0.88	0.12	0.99	0.25	
Red/green gratings	7.88	6.46	1.21	1.10	0.46	0.95	0.21	
Blue/yellow gratings	7.79	6.37	1.14	1.10	0.59	1.00	0.24	
Flickering discs	7.72	6.32	1.26	0.83	0.22	0.90	0.23	

grating conditions, and slightly below 1 (0.89) for our flickering discs. The exponent values ( $p$  and  $q$ ) for the red/green gratings are rather different from those of the other conditions. In previous work (Meese et al., 2006) the second stage gain control nonlinearity (Equation 4) does typically have quite substantial exponent values, which balance the relatively mild nonlinearity at the first stage, and produce a compressive transducer that results in contrast masking (the handle of the dipper). The high value of  $p$  (19.42) is therefore likely to be compensating for the low value of  $m$  (1.03) at stage one, but may well represent the combination of several successive stages of nonlinearity, and perhaps also other phenomena such as uncertainty (Pelli, 1985).

We additionally implemented a hierarchical Bayesian version of the model in order to estimate full posterior parameter distributions. This model was fit simultaneously to the full trial-by-trial data from all participants who participated in a given experiment (i.e. the model was fitted separately to each of the four dipper data sets). Figure 7a-d summarises the model behaviour, which displays the same pattern of dipper functions as we found empirically. Mean posterior parameter values are given in the lower rows of Table 1. These correspond quite closely to the parameters from the simplex fitting and the original Meese et al. (2006) parameters (reproduced in the first row of the table for reference). The panels in the right and upper margins of Figure 7 show posterior distributions for each model parameter. Note in particular that the posterior distribution for the weight of interocular suppression ( $\omega$ , top right panel) overlaps 1 for each experiment, consistent with the strong dichoptic masking observed in the threshold data. We note that the distribution for  $\omega$  in the flickering disc condition is somewhat lower than for the other three data sets. However this difference is not meaningful according to criteria such as comparing the 95% credible intervals of the distribution to a value of  $\omega = 1$ .

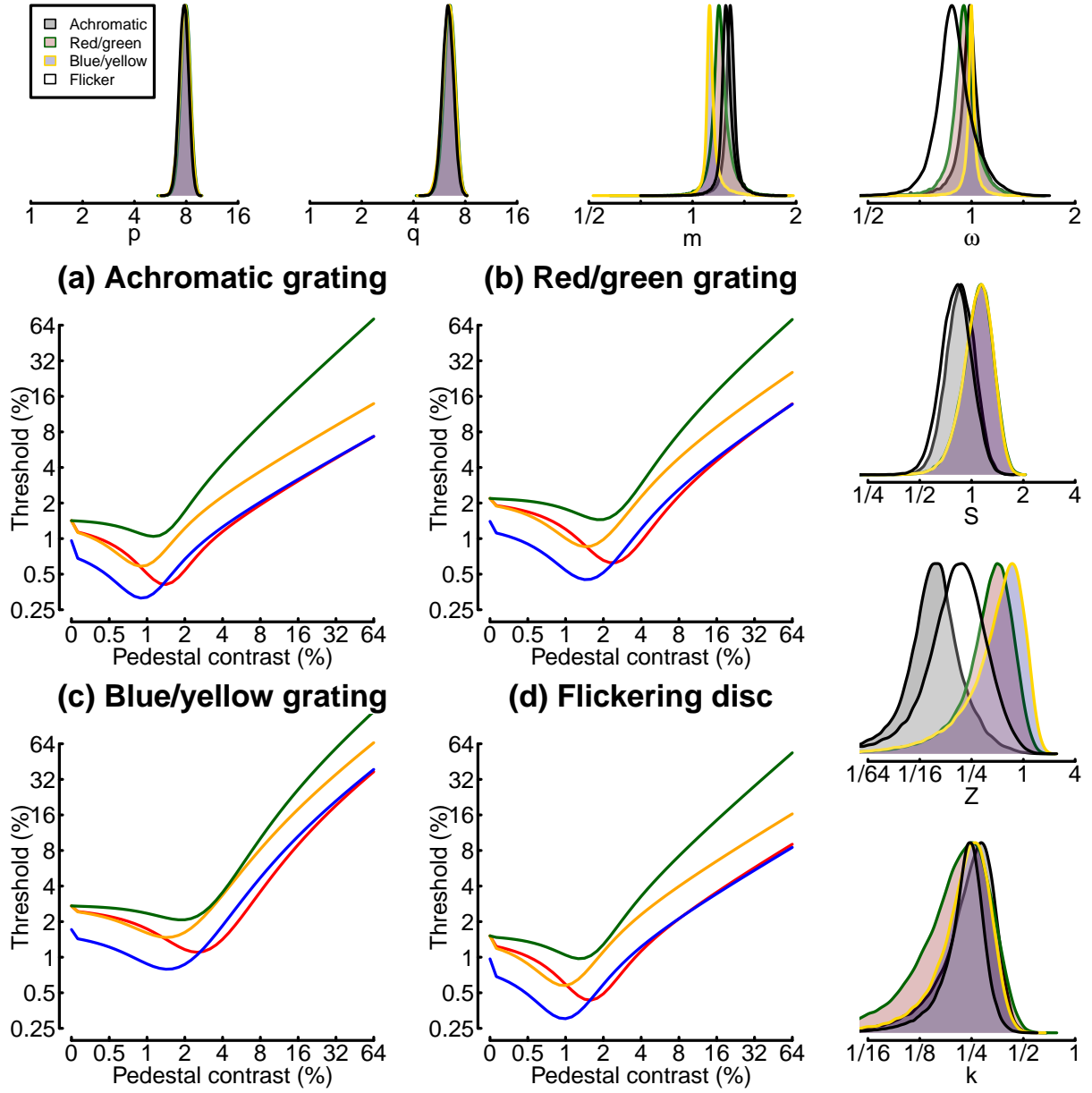


Figure 7: Model predictions (a-d) and posterior parameters (top and right margin plots) for the hierarchical Bayesian model. Thick lines in panels (a-d) show curves generated from the mean posterior parameter estimates. The probability density functions in the margin plots are peak-normalized, and shown for each of the four data sets in different colours (see legend in upper left plot). Distributions were generated from 1000000 samples per data set, using a Markov Chain Monte Carlo sampling algorithm. Note the logarithmic x-axis for all posterior plots.

Finally we fitted an extended model, that included interocular suppression between the different pathways, to the data of Experiment 2. Figure 8 shows the model curves (panels a-c), which correspond closely to the data in Figure 5. One striking discrepancy is that the model predicts a region of negative  $d'$  for the case where the target and dichoptic mask have the same chromaticity (see the curved regions below 50% correct). This feature was not generally present in our data, though it can be observed for one participant in Figure A2b,c. Our previous work (Baker et al., 2013) has found evidence for this phenomenon (termed ‘swan

functions'), but it generally requires very high mask contrasts to be measurable empirically, and there may also be some individual differences; both factors might explain its absence here.

Figure 8d,e shows the model estimates for the weight of interocular suppression in each combination of target and mask chromaticity. Although the weights might appear to indicate that (excluding within-mechanism effects) suppression is strongest between blue/yellow targets and achromatic masks (a weight of 0.56), this is somewhat complicated by the differences in sensitivity and mask contrast between conditions. We selected mask contrasts that were approximately equal multiples ( $16\times$ ) of their own detection thresholds, based on the data of Experiment 1 (Figure 4a-c). However the modelling expressed these contrasts as a percentage of the maximum contrast that it was possible to display on the monitor (see Procedures). The values in brackets in Figure 8d indicate the overall suppressive term on the denominator, calculated by multiplying the mask contrast by the fitted weight. These values indicate that the strongest suppression is instead between red/green targets and blue/yellow masks, consistent with the threshold elevation observed empirically (Figure 5g). In principle it would be possible to repeat this modelling using either cone contrast values, or contrast expressed relative to detection threshold, which would require re-fitting the model.

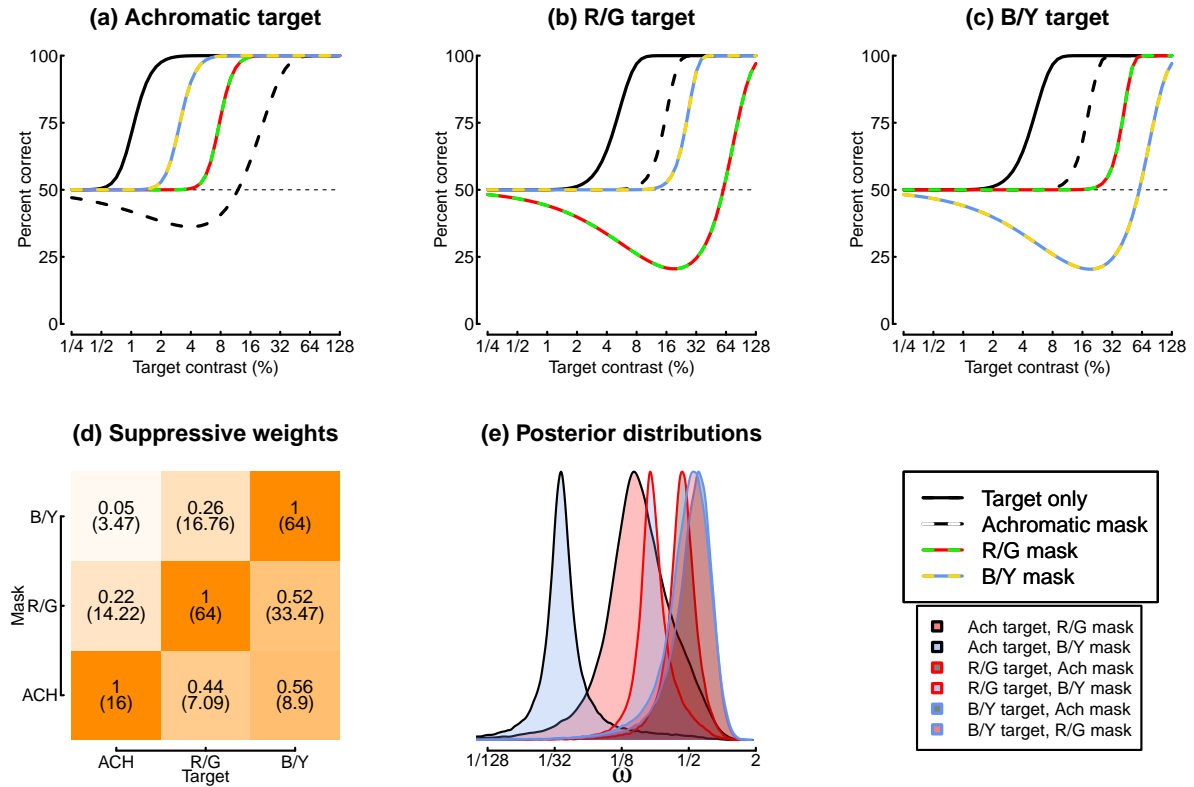


Figure 8: Summary of the model fit to the data of Experiment 2. Panels (a-c) show model psychometric functions in the same format as those in Figure 5. Panel (d) shows the fitted suppressive weights ( $\omega$  values), with values in brackets indicating the combined suppressive denominator term calculated by multiplying together the suppressive weight and the dichoptic mask contrast. Panel (e) shows the posterior distributions for each weight parameter. The lower right plot gives the legends for the upper row (top) and the posterior distributions in panel (e).

## 5 Discussion

Across three psychophysical experiments, we have demonstrated that:

- Binocular combination of isoluminant chromatic stimuli is similar to that for achromatic stimuli
- Interocular suppression is strongest within a post-retinal pathway, and weakest between achromatic and chromatic pathways
- Binocular combination occurs similarly for spatial and temporal luminance modulations

We now discuss the relationship to previous work, and consider the likely physiological substrates of these effects.

## 5.1 Summation at threshold

Estimates of the binocular summation ratio at threshold fell within the range  $\sqrt{2}$  to 2 for all four data sets. Consistent with previous reports (Simmons, 2005), summation was slightly higher for chromatic versus achromatic stimuli, but fell short of perfect linear summation (a ratio of 2). In our model, summation at threshold is determined by the exponent at the first gain control stage (the  $m$  parameter in Equations 1 & 2). The summation ratio can be approximated by  $2^{1/m}$  (see Figure 9 of Baker et al., 2012), such that an exponent of  $m = 1$  produces a ratio of 2, and an exponent of  $m = 2$  produces a ratio of  $\sqrt{2}$ . Both the simplex and Bayesian model fits (see Table 1) generated parameter estimates in the range  $1 < m < 1.4$ , consistent with the high levels of summation observed empirically.

## 5.2 Interocular suppression

The weight of interocular suppression is represented by the model parameter,  $\omega$  (Equations 1 & 2). Posterior distributions of this parameter overlapped 1 for all of our dipper function data sets (see top right panel in Figure 7), indicating strong interocular suppression regardless of post-retinal pathway. The flickering disc stimuli produced a slightly lower suppression estimate, with  $\omega \sim 0.9$ . This may indicate slightly weaker suppression between the eyes for temporal modulations, but it is much less extreme than our recent estimates using EEG and matching paradigms (Segala et al., 2023). One difference between these studies is the luminance of the background, which was set to black for the experiments of Segala et al. (2023), but here was set at the mean luminance. Other studies using static luminance increments have also reported differences in the character of binocular combination that are attributable to background luminance (Baker et al., 2012), so this may explain the differences between studies. We also note that the weaker interocular suppression for flickering disc stimuli appears to be largely due to one participant (see Figure A3c), so individual differences might also play a role. Future work should manipulate the background luminance systematically to better understand how this modulates interocular suppression.

## 5.3 Psychometric slopes

Previous work has demonstrated that the slope of the psychometric function in 2AFC tasks can distinguish different types of masking, although it is much less widely reported than threshold measures. In particular, pedestal masking linearizes the slope (Foley & Legge, 1981; Meese et al., 2006), and within-channel dichoptic masking produces very steep slopes (Baker et al., 2013; Meese et al., 2006). We replicate both of these effects here, and show that they extend to isoluminant chromatic pathways (see Figures 4d-f, 5). We additionally show that dichoptic masking between different pathways does not produce unusually steep slopes (see Figure 5). It therefore more closely resembles other types of masking between visual channels, such as cross-orientation masking (Meese & Baker, 2009), surround masking (Yu et al., 2002), and masking from broadband noise (Baker & Meese, 2012; Lu & Doshier, 2008), which also do not impact psychometric slopes.

## 5.4 Physiological substrates

Recent evidence indicates that the physiological substrate of interocular suppression may be neurons in layer 4 of primary visual cortex (Dougherty et al., 2019). Most cells in this layer are monocularly excitable, in that their responses increase only by stimulation of their preferred eye. However, simultaneous stimulation of the non-preferred eye can modulate the response, usually in an inhibitory fashion, exactly as proposed at stage 1 of the two-stage model (Equations 1 & 2). In terms of perception, one consequence of this early suppression is to achieve ‘ocularly invariance’, whereby the perceived contrast of a stimulus viewed by one

eye is equivalent to that of the same stimulus viewed by both eyes (Baker et al., 2007). Similar processes of invariance have also been reported using fMRI (Moradi & Heeger, 2009) and steady-state EEG (Baker & Wade, 2017).

In V1, chromatic stimuli are processed in ‘blob’ regions that are largely monocular as they fall within ocular dominance columns (Livingstone & Hubel, 1984). Since our psychophysical results indicate that interocular suppression is equally strong within chromatic and achromatic pathways, it may be that blob regions include interocular suppression. Presumably binocular summation of signals occurs at a later stage of processing for chromatic stimuli, but this does not lead to different perceptual outcomes. This physiological arrangement might also be responsible for the pattern of between-pathway interocular suppression (see Figure 5g). If blobs process signals from both the L-M and S-(L+M) pathways, their close physical proximity might explain the stronger suppression between the two chromatic pathways than between the chromatic and (more distant) achromatic pathways.

## 6 Conclusions

Here we provide estimates of interocular suppression within and between the three primary post-retinal visual pathways. These results show that binocular signal combination is similar within each pathway, but that interocular suppression is relatively weak between pathways. Our findings could be applied when building models to predict perception of binocular images and movies, for example those generated by virtual and augmented reality systems, or in 3D cinema and television.

## 7 Acknowledgements

Supported by BBSRC grant BB/V007580/1 awarded to DHB and ARW.

## 8 References

- Anstis, S., & Ho, A. (1998). Nonlinear combination of luminance excursions during flicker, simultaneous contrast, afterimages and binocular fusion. *Vision Res*, 38(4), 523–539. [https://doi.org/10.1016/s0042-6989\(97\)00167-3](https://doi.org/10.1016/s0042-6989(97)00167-3)
- Baker, D. H., Lygo, F. A., Meese, T. S., & Georgeson, M. A. (2018). Binocular summation revisited: Beyond  $\sqrt{2}$ . *Psychol Bull*, 144(11), 1186–1199. <https://doi.org/10.1037/bul0000163>
- Baker, D. H., & Meese, T. S. (2007). Binocular contrast interactions: Dichoptic masking is not a single process. *Vision Res*, 47(24), 3096–3107. <https://doi.org/10.1016/j.visres.2007.08.013>
- Baker, D. H., & Meese, T. S. (2012). Zero-dimensional noise: The best mask you never saw. *J Vis*, 12(10). <https://doi.org/10.1167/12.10.20>
- Baker, D. H., Meese, T. S., & Georgeson, M. A. (2007). Binocular interaction: Contrast matching and contrast discrimination are predicted by the same model. *Spat Vis*, 20(5), 397–413. <https://doi.org/10.1163/156856807781503622>
- Baker, D. H., Meese, T. S., & Georgeson, M. A. (2013). Paradoxical psychometric functions ("swan functions") are explained by dilution masking in four stimulus dimensions. *iPerception*, 4(1), 17–35. <https://doi.org/10.1068/i0552>
- Baker, D. H., & Wade, A. R. (2017). Evidence for an optimal algorithm underlying signal combination in human visual cortex. *Cereb Cortex*, 27(1), 254–264. <https://doi.org/10.1093/cercor/bhw395>
- Baker, D. H., Wallis, S. A., Georgeson, M. A., & Meese, T. S. (2012). Nonlinearities in the binocular combination of luminance and contrast. *Vision Res*, 56, 1–9. <https://doi.org/10.1016/j.visres.2012.01.008>
- Campbell, F. W., & Green, D. G. (1965). Monocular versus binocular visual acuity. *Nature*, 208(5006), 191–192. <https://doi.org/10.1038/208191a0>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32. <https://doi.org/10.18637/jss.v076.i01>
- Chen, C., Foley, J. M., & Brainard, D. H. (2000). Detection of chromoluminance patterns on chromoluminance



- pedestals i: Threshold measurements. *Vision Res*, 40(7), 773–788. [https://doi.org/10.1016/s0042-6989\(99\)00227-8](https://doi.org/10.1016/s0042-6989(99)00227-8)
- Dougherty, K., Cox, M. A., Westerberg, J. A., & Maier, A. (2019). Binocular modulation of monocular V1 neurons. *Curr Biol*, 29(3), 381–391.e4. <https://doi.org/10.1016/j.cub.2018.12.004>
- Edwards, D. P., Purpura, K. P., & Kaplan, E. (1995). Contrast sensitivity and spatial frequency response of primate cortical neurons in and around the cytochrome oxidase blobs. *Vision Res*, 35(11), 1501–1523. [https://doi.org/10.1016/0042-6989\(94\)00253-i](https://doi.org/10.1016/0042-6989(94)00253-i)
- Foley, J. M., & Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Res*, 21(7), 1041–1053. [https://doi.org/10.1016/0042-6989\(81\)90009-2](https://doi.org/10.1016/0042-6989(81)90009-2)
- Georgeson, M. A., Wallis, S. A., Meese, T. S., & Baker, D. H. (2016). Contrast and lustre: A model that accounts for eleven different forms of contrast discrimination in binocular vision. *Vision Res*, 129, 98–118. <https://doi.org/10.1016/j.visres.2016.08.001>
- Horton, J. C., & Hubel, D. H. (1981). Regular patchy distribution of cytochrome oxidase staining in primary visual cortex of macaque monkey. *Nature*, 292(5825), 762–764. <https://doi.org/10.1038/292762a0>
- Kim, Y. J., Gheiratmand, M., & Mullen, K. T. (2013). Cross-orientation masking in human color vision: Application of a two-stage model to assess dichoptic and monocular sources of suppression. *J Vis*, 13(6), 15. <https://doi.org/10.1167/13.6.15>
- Kim, Y. J., & Mullen, K. T. (2015). The dynamics of cross-orientation masking at monocular and interocular sites. *Vision Res*, 116(Pt A), 80–91. <https://doi.org/10.1016/j.visres.2015.09.008>
- Kingdom, F. A. A., & Libenson, L. (2015). Dichoptic color saturation mixture: Binocular luminance contrast promotes perceptual averaging. *J Vis*, 15(5), 2. <https://doi.org/10.1167/15.5.2>
- Legge, G. E. (1979). Spatial frequency masking in human vision: Binocular interactions. *J Opt Soc Am*, 69(6), 838–847. <https://doi.org/10.1364/josa.69.000838>
- Legge, G. E. (1984). Binocular contrast summation–II. Quadratic summation. *Vision Res*, 24(4), 385–394. [https://doi.org/10.1016/0042-6989\(84\)90064-6](https://doi.org/10.1016/0042-6989(84)90064-6)
- Levelt, W. J. (1965). Binocular brightness averaging and contour information. *Br J Psychol*, 56, 1–13. <https://doi.org/10.1111/j.2044-8295.1965.tb00939.x>
- Livingstone, M. S., & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *J Neurosci*, 4(1), 309–356. <https://doi.org/10.1523/JNEUROSCI.04-01-00309.1984>
- Lu, Z.-L., & Doshier, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychol Rev*, 115(1), 44–82. <https://doi.org/10.1037/0033-295X.115.1.44>
- Maehara, G., & Goryo, K. (2005). Binocular, monocular and dichoptic pattern masking. *Optical Review*, 12(2), 76–82. <https://doi.org/10.1007/PL00021542>
- Medina, J. M., & Mullen, K. T. (2009). Cross-orientation masking in human color vision. *J Vis*, 9(3), 20.1–16. <https://doi.org/10.1167/9.3.20>
- Meese, T. S., & Baker, D. H. (2009). Cross-orientation masking is speed invariant between ocular pathways but speed dependent within them. *J Vis*, 9(5), 2.1–15. <https://doi.org/10.1167/9.5.2>
- Meese, T. S., Georgeson, M. A., & Baker, D. H. (2006). Binocular contrast vision at and above threshold. *J Vis*, 6(11), 1224–1243. <https://doi.org/10.1167/6.11.7>
- Moradi, F., & Heeger, D. J. (2009). Inter-ocular contrast normalization in human visual cortex. *J Vis*, 9(3), 13.1–22. <https://doi.org/10.1167/9.3.13>
- Mullen, K. T., Kim, Y. J., & Gheiratmand, M. (2014). Contrast normalization in colour vision: The effect of luminance contrast on colour contrast detection. *Sci Rep*, 4, 7350. <https://doi.org/10.1038/srep07350>
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *J Opt Soc Am A*, 2(9), 1508–1532. <https://doi.org/10.1364/josaa.2.001508>
- Schütt, H. H., Harmeling, S., Macke, J. H., & Wichmann, F. A. (2016). Painfree and accurate bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Res*, 122, 105–123. <https://doi.org/10.1016/j.visres.2016.02.002>
- Segala, F. G., Bruno, A., Martin, J. T., Aung, M. T., Wade, A. R., & Baker, D. H. (2023). Different rules for binocular combination of luminance flicker in cortical and subcortical pathways. *eLife*, 12, RP87048. <https://doi.org/10.7554/eLife.87048>
- Simmons, D. R. (2005). The binocular combination of chromatic contrast. *Perception*, 34(8), 1035–1042. <https://doi.org/10.1068/p5279>

- Stockman, A., & Sharpe, L. T. (2000). The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Res*, *40*(13), 1711–1737. [https://doi.org/10.1016/s0042-6989\(00\)00021-3](https://doi.org/10.1016/s0042-6989(00)00021-3)
- Tootell, R. B., Silverman, M. S., Hamilton, S. L., De Valois, R. L., & Switkes, E. (1988). Functional anatomy of macaque striate cortex. III. color. *J Neurosci*, *8*(5), 1569–1593. <https://doi.org/10.1523/JNEUROSCI.08-05-01569.1988>
- Wallis, S. A., Baker, D. H., Meese, T. S., & Georgeson, M. A. (2013). The slope of the psychometric function and non-stationarity of thresholds in spatiotemporal contrast vision. *Vision Res*, *76*, 1–10. <https://doi.org/10.1016/j.visres.2012.09.019>
- Yu, C., Klein, S. A., & Levi, D. M. (2002). Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanisms. *J Vis*, *2*(3), 243–255. <https://doi.org/10.1167/2.3.4>

## 9 Appendices

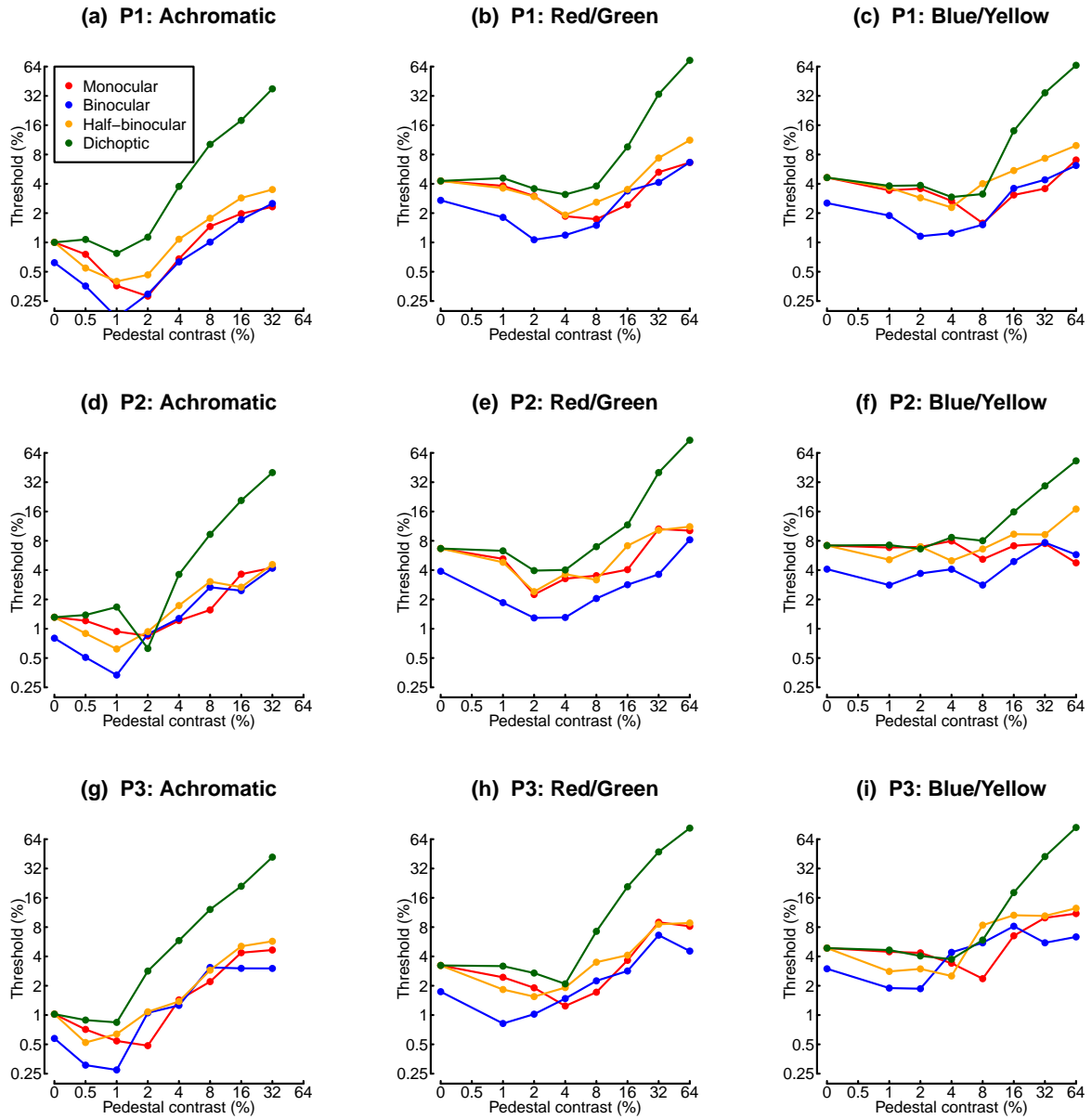


Figure A1: Individual participant data from Experiment 1.

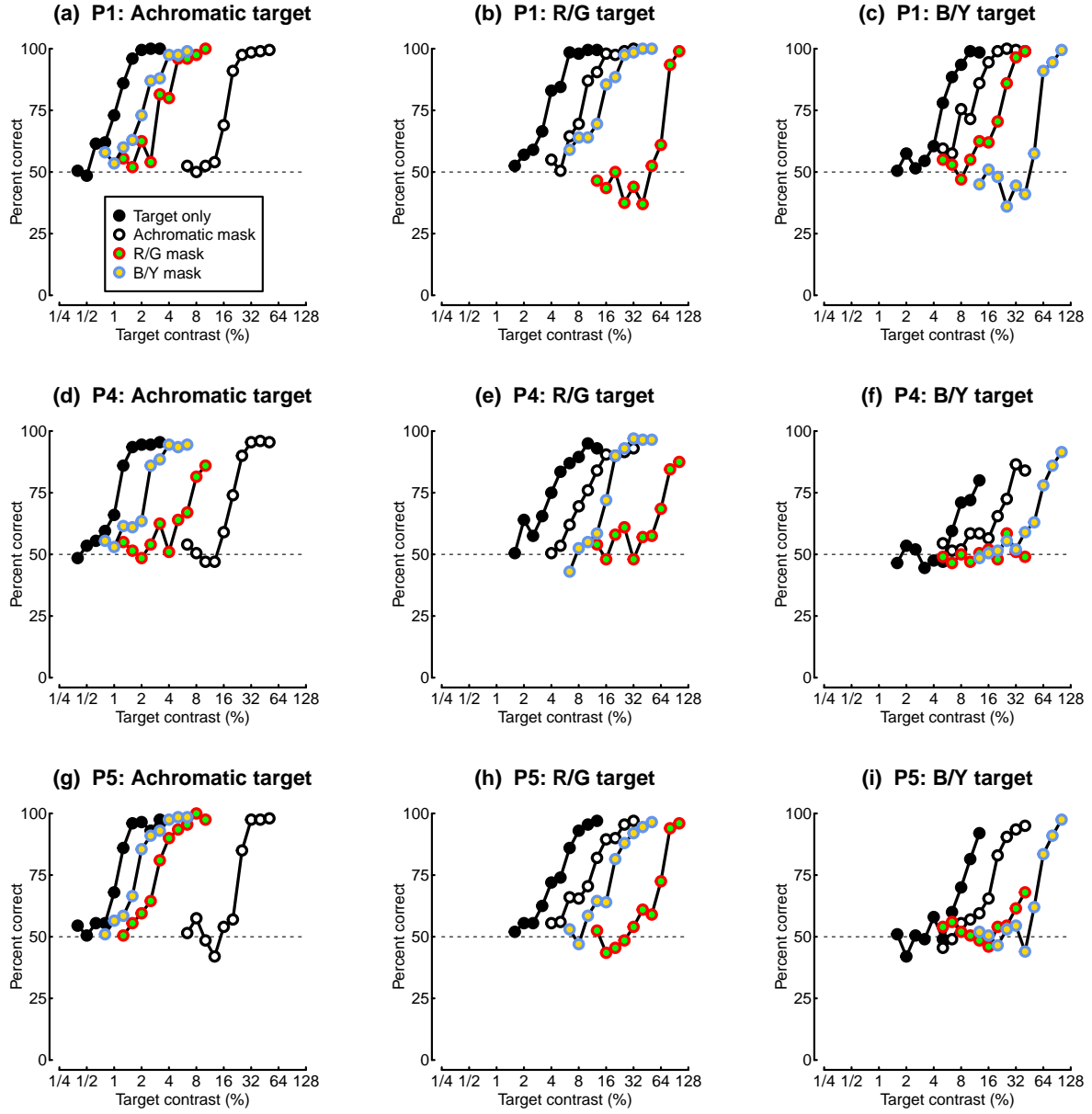


Figure A2: Individual participant data from Experiment 2.

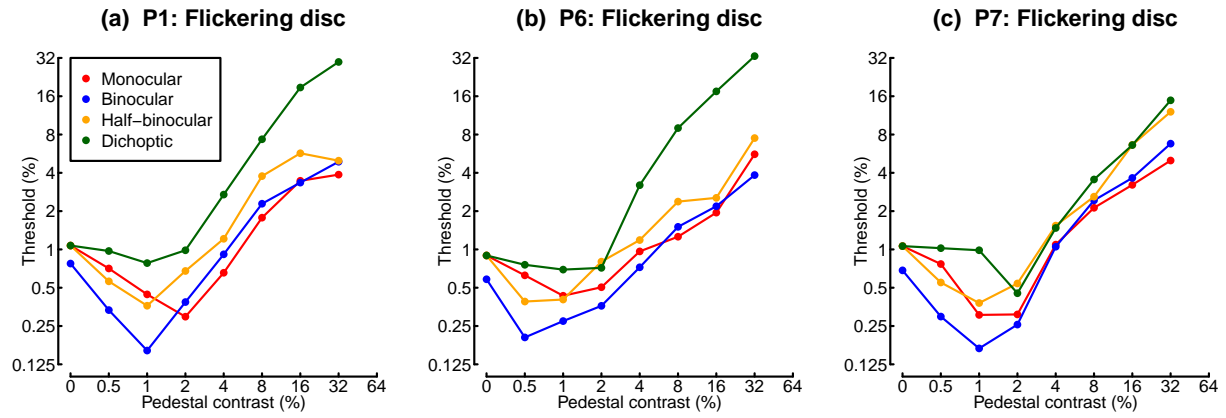


Figure A3: Individual participant data from Experiment 3.