Interocular difference thresholds are mediated by binocular differencing, not summing, channels

Frederick A. A. Kingdom

McGill Vision Research, Department of Ophthalmology, Montréal General Hospital, Montréal, Canada 🏠 ✉

Nour M. Seulami

McGill Vision Research, Department of Ophthalmology, Montréal General Hospital, Montréal, Canada ✉

Ben J. Jennings

Centre for Cognitive Neuroscience, Division of Psychology, College of Health and Life Science, Brunel University London, London, UK 🏠 ✉

Mark A. Georgeson

School of Life & Health Sciences, Aston University, Birmingham, UK 🏠 ✉

Patterns in the two eyes' views that are not identical in hue or contrast often elicit an impression of luster, providing a cue for discriminating them from perfectly matched patterns. Here we attempt to determine the mechanisms for detecting interocular differences in luminance contrast, in particular in relation to the possible contributions of binocular differencing and binocular summing channels. Test patterns were horizontally oriented multi-spatial-frequency luminance-grating patterns subject to variable amounts of interocular difference in grating phase, resulting in varying degrees of *local* interocular contrast difference. Two types of experiment were conducted. In the first, subjects discriminated between a pedestal with an interocular difference that ranged upward from zero (i.e., binocularly correlated) and a test pattern that contained a bigger interocular difference. In the second type of experiment, subjects discriminated between a pedestal with an interocular difference that ranged downward from a maximum (i.e., binocularly anticorrelated) and a test pattern that contained smaller interocular difference. The two types of task could be mediated by a binocular differencing and a binocular summing channel, respectively. However, we found that the results from both experiments were well described by a simpler model in which a single, linear binocular differencing channel is followed by a standard nonlinear transducer that is expansive for small signals but strongly compressive for large ones. Possible reasons for the lack of involvement of a binocular summing channel are discussed in the context of a model that incorporates the responses of both monocular and binocular channels.

Introduction

Two spatially separated eyes with overlapping visual fields form the basis of binocular vision, an arrangement that benefits the user with a wider field of view, stereopsis, binocular summation, and binocular difference detection. The last of these, binocular difference detection, is a subject of increasing interest (Julesz & Tyler, 1976; Tyler & Julesz, 1976; Cohn, Leong, & Lasley, 1981; Julesz, 1986; Cormack, Stevenson, & Schor, 1991; Stevenson, Cormack, Schor, & Tyler, 1992; Formankiewicz & Mollon, 2009; Yoonessi & Kingdom, 2009; Malkoc & Kingdom, 2012; Georgeson, Wallis, Meese, & Baker, 2016; Jennings & Kingdom, 2016; Kingdom, Jennings, & Georgeson, 2018; Reynaud & Hess, 2018). Binocular differences have been termed *interocular (de)-correlations* (Cormack et al., 1991; Stevenson et al., 1992; Reynaud & Hess, 2018), *dichoptic differences* (e.g., Yoonessi & Kingdom, 2009; Malkoc & Kingdom, 2012) *binocular luminance disparities* (Formankiewicz & Mollon, 2009), and simply *interocular differences*, the term we will employ here. An interocular difference in contrast or hue can generate an
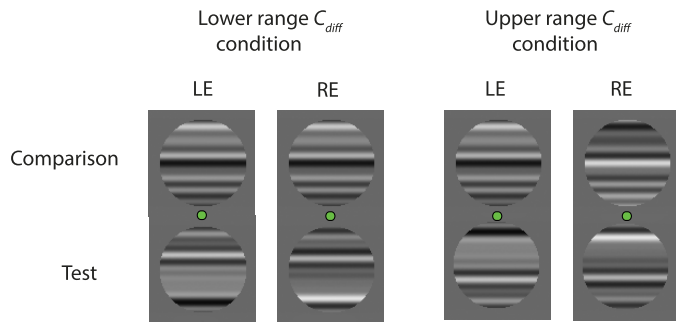
Figure 1. Example stimuli used in Experiment 1. Left: example of lower-range C_{diff} condition with comparison stimulus identical in the two eyes—i.e., binocularly correlated—and test stimulus less correlated. Right: example of upper-range C_{diff} condition with comparison stimulus anticorrelated in the two eyes and test stimulus more correlated. LE = left eye; RE = right eye.

impression of luster, a cue that has been argued to enable detection of interocular differences (Forman-kiewicz & Mollon, 2009; Yoonessi & Kingdom, 2009; Malkoc & Kingdom, 2012; Jennings & Kingdom, 2016; Kingdom et al., 2018). Recent studies have suggested models for interocular difference detection based on luster (Georgeson et al., 2016; Jennings & Kingdom, 2016) and have furthermore demonstrated that interocular difference detection is an adaptable dimension of vision (Kingdom et al., 2018).

In this article, we probe the mechanisms involved in interocular difference detection, specifically to assess the involvement of binocular differencing (B−) and binocular summing (B+) channels. The involvement of B− channels in binocular vision is evident from studies of contrast detection (Cohn et al., 1981), motion perception (May, Zhaoping, & Hibbard, 2012; see also Kingdom, 2012), orientation perception (May & Zhaoping, 2016), stereopsis (Goncalves & Welchman, 2017; Kingdom, Yared, Hibbard, & May, in press), binocular rivalry (Said & Heeger, 2013), visual-evoked potentials (Katyal, Vergeer, He, He, & Engel, 2018), and interocular difference detection (Kingdom et al., 2018). Involvement of B+ channels has also emerged from many of these studies, but its main support comes from the plethora of studies demonstrating substantial improvements in thresholds for detecting stimuli when viewed by both eyes compared to one (see recent review and metanalysis by Baker, Lygo, Meese, & Georgeson, 2018), as well as from studies modeling the appearance of dichoptic mixtures of stimuli differing in luminance or color contrast (Hovis, 1989; Baker, Wallis, Georgeson, & Meese, 2012; Kingdom & Libenson, 2015).

Intuitively, one would expect the B− channel to mediate the detection of interocular differences, so why the potential involvement of the B+ channel? This question lies at the heart of the rationale for the present study. On the left of Figure 1 are shown

dichoptic pairs of the grating stimuli employed in the present study, details of which are provided later. In these stimuli the interocular differences are introduced via spatial phase differences between the component sine-wave gratings in the dichoptic pairs, within the range 0°–180°. However, in keeping with our previous study showing that adaptation of interocular differences was best understood if interocular difference was expressed in terms of root mean square (RMS) local contrast difference C_{diff} (Kingdom et al., 2018), we use this as our measure here. For the situation in which the contrasts in the two eyes are the same, C_{diff} is given by

$$C_{\mathrm{diff}} = C\sqrt{2(1 - \cos\phi)}, \quad (1)$$

where C is the RMS contrast of the image for each eye, ϕ is the interocular difference in grating phase, and contrast is the same for all the component spatial frequencies of the image.

The left of Figure 1 shows one of our conditions. It comprises two dichoptic pairs, the upper one perfectly interocularly correlated—that is, with a ϕ and C_{diff} of zero—and the other with a nonzero ϕ and hence positive C_{diff}. On the right is shown a second condition, where the upper pair is interocularly anticorrelated—that is, of opposite luminance polarity between the eyes—produced by setting ϕ to 180° and resulting in a maximum C_{diff}. The other pair has a smaller ϕ and hence smaller C_{diff}. In the experiments to be described, our observers were required to discriminate between pairs of stimuli with different C_{diff}, to determine their just-noticeable differences (JNDs). We did this for both the lower range of C_{diff}, as exemplified by the left-hand figure, and the upper range, as exemplified by the right-hand figure.

Visual mechanisms are often compressive in their response to the magnitude of the dimension to which they are sensitive, so for the lower range of C_{diff} we would expect JNDs to be mediated by the B− channel, as it would signal the difference between zero and some positive value, using the early, noncompressive part of its response range. On the other hand, in the upper range of C_{diff} the JNDs would be best served not by the B− channel, as it would be operating within its compressive response range, but instead by the B+ channel, again because it would be signaling the difference between zero and some positive value. Figure 2 helps to reinforce the point by showing how the binocular contrast difference (in red) and the binocular sum (in green) of a single dichoptic pair change as a function of ϕ, where the binocular sum C_{sum} is given by

$$C_{\mathrm{sum}} = C\sqrt{2(1 + \cos\phi)}. \quad (2)$$

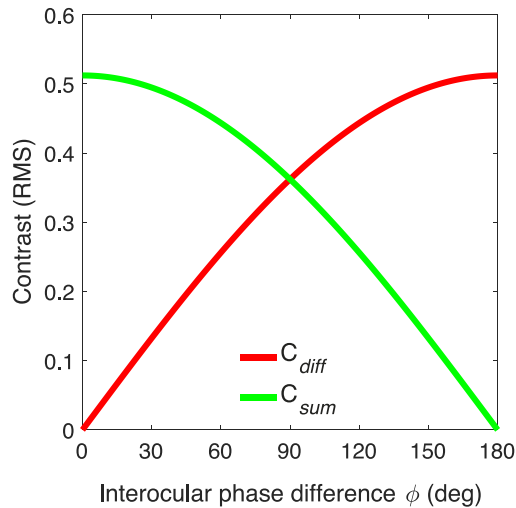As Figure 2 shows, C_{diff} increases and saturates at large phase disparities, whereas C_{sum} does the

Figure 2. Interocular contrast difference C_{diff} (red) and sum C_{sum} (green) expressed in root mean square contrast as a function of the interocular phase difference ϕ for the stimuli as exemplified in Figure 1.

General methods

Observers

Seven observers took part in the experiments. Three were authors; however, one of those authors was, at the time of testing, unaware of the purpose of the experiment. The remaining four observers were all undergraduate volunteers who were unaware of the experimental purpose. All observers had normal or corrected-to-normal visual acuity. Prior to experimental testing, informed consent was obtained from each observer. All experiments were conducted in accordance with the Declaration of Helsinki and the Research Institute of the McGill University Health Centre (RI-MUHC) Ethics Board. Observer initials on graphs have been anonymized in accordance with requirements of the Ethics Board. Observers 1–6 participated in Experiment 1, and Observers 1, 2, 3, and 7 in Experiment 2.

Stimulus display

All experiments were conducted using a Dell Precision T1650 PC with a ViSaGe graphics card (Cambridge Research Systems, Rochester, UK). The visual stimuli were displayed on a gamma-corrected Sony Trinitron Multiscan F500 flat-screen CRT monitor. Stimulus generation and experimental control use custom software written in C. Participants viewed the dichoptic pairs through a custom-built eight-mirror stereoscope with an aperture of $10° \times 10°$ and a viewing distance along the light path of 55 cm. During the experiments, observers were seated in a darkened room and their responses were recorded via a keypad.

Stimuli

The stimulus images for all experiments are illustrated in Figure 1a. They were dichoptic pairs of circular patches, each with a diameter of 4.35°. The horizontal separation of the two members of each dichoptic pair on the monitor was adjusted so that they appeared fused in the center of the aperture. The two members of each two-alternative forced-choice pair were presented together one above the other, separated vertically by 5.8° center to center, above and below a small green spot contained within a black fixation circle 0.27° in diameter which helped maintain vergence. Each patch comprised eight sine-wave luminance gratings of equal contrast, with spatial frequencies (SFs) of 1, 2, 3, 4, 5, 6, 7, and 8 c/patch, corresponding to spatial frequencies ranging from 0.23 to 1.84 c/°. The base

opposite. Note that the saturated parts of these curves are a physical property of the way that the sum and difference signals vary with phase disparity, a property that would only be exacerbated by an internal compressive transducer. The main point, however, is that for interocular pairs that fall within the range of $\phi = 0°$ to $90°$, the B– channel (responding to C_{diff}) would be expected to be most differentially responsive, whereas for pairs that fall within the range of $\phi = 90°$ to $180°$, the B+ channel (responding to C_{sum}) would be expected to be most differentially responsive. Our main aim is to test this prediction by comparing performance found for the two ranges of ϕ illustrated in Figure 1.

Why do we manipulate C_{diff} by varying the interocular phase difference ϕ between the dichoptic images rather than by varying the relative contrasts of the two monocular gratings? Our reasons have been detailed elsewhere (Kingdom et al., 2018), but in brief the use of phase difference is first because it minimizes the possibility that global contrast can be used as a cue to the presence of an interocular difference (because RMS contrast is the same in both eyes and the same for all ϕ) and second because of the simple mathematical relationship between interocular phase difference and local interocular contrast difference, as in Equation 1. Finally, our use of horizontally oriented gratings minimizes stereo-depth cues to the stimulus containing the interocular difference, because horizontally oriented gratings have only vertical disparities, and these appear to play no role in depth perception, at least in central vision.

spatial phase ϕ_0 of each grating component was randomized across SF, but the magnitude of phase disparity ϕ was the same for each SF, with the sign of this disparity randomized across SF. Thus the component phase for the left eye was $(\phi_0 + a.\phi/2)$, and for the right eye $(\phi_0 - a.\phi/2)$, where a was randomly 1 or -1 across SF. The randomization of ϕ_0 and a did introduce random variations in the waveform structure (see Figure 1) but did not perturb the value of C_{diff}. One member of each two-alternative forced-choice pair comprised a fixed, or pedestal, level of C_{diff}, and the other a pedestal plus or minus a variable ΔC_{diff}. In the experiment exploring the lower range of C_{diff}, the pedestal C_{diff} involved ϕ ranging from 0° to 90°, with ΔC_{diff} an increment. In the experiments exploring the upper range of C_{diff}, the fixed level of C_{diff} involved ϕ ranging from 180° to 90°, with ΔC_{diff} a decrement. We will refer to the two types of experiment as the lower- and upper-range C_{diff} or ϕ experiments.

Procedure

We employed a conventional two-alternative forced-choice method in conjunction with a staircase procedure that adjusted C_{diff} according to previous responses. The base phase ϕ_0 of every SF component was randomized afresh for every stimulus presentation. Stimulus exposure duration was 500 ms. Each stimulus presentation was initiated by a button press in response to the previous stimulus, enabling observer control over the trial sequence. Each stimulus was preceded by a spatially uniform blank field at mean luminance for 500 ms and was followed by a blank field for 250 ms and then a feedback signal in which the central green fixation dot turned red for 100 ms if the response was incorrect. After 50 trials the session was terminated. In the experiments exploring the lower range of C_{diff}, the task on each trial was to identify the position (upper or lower) of the patch containing the bigger C_{diff}—observers were instructed to "select the stimulus with the most luster." In the experiments exploring the upper range of C_{diff}, the task was to identify the position of the smaller C_{diff}—observers were instructed to "choose the stimulus with the least luster." The different instructions for the two tasks ensured that the two tasks were comparable in terms of what constituted the target stimulus—that is, the one that was varied during the staircase procedure: the more lustrous stimulus for the first task and the less lustrous for the second. The initial difference in C_{diff} between the members of the forced-choice pair, ΔC_{diff}, was randomly selected from a range whose average was approximately double the expected threshold ΔC_{diff} as determined in pilot runs. A 3-up-1-down staircase method was used in which ΔC_{diff} either

increased or decreased proportionately on each trial by a factor of 2.5 for the first five trials and a factor of 1.3 thereafter. Correct and incorrect trial sequences resulted in, respectively, decreases and increases in the magnitude of ΔC_{diff}, with the sign of ΔC_{diff} always being positive for the lower and negative for the upper range. There were between five and 10 sessions for each condition, resulting in a total of between 250 and 500 trials per condition. Condition order was randomized.

Analysis

Psychometric functions of proportion correct against ΔC_{diff} were fitted with Quick functions using a maximum-likelihood criterion, using routines customized from the Palamedes toolbox (Prins & Kingdom, 2018). Threshold ΔC_{diff} at the 75% correct level (where performance $d' = 0.954$ is close to 1) and associated bootstrap errors were estimated from the fits.

Experiment 1: Threshold ΔC_{diff} with correlated and anticorrelated comparisons

In the first experiment we compared JNDs, expressed as ΔC_{diff}, for the lower and upper extremes of the C_{diff} range—that is, using comparison stimuli with $\phi = 0°$ and $\phi = 180°$, respectively, as shown in Figure 1. Results for six observers are shown in Figure 3.

Although for the 180° comparison condition ΔC_{diff} was a decrement, the absolute value is given in the figure to allow a direct comparison with the increment ΔC_{diff} measured in the 0° comparison condition. The results show that for all observers, ΔC_{diff} was significantly lower for the 0° condition than the 180° condition. The geometric mean ratio of ΔC_{diff} values for the two conditions, averaged across the six observers, is 4.37. This shows that ΔC_{diff} detection is markedly asymmetric, in that detecting a change in C_{diff} between two image pairs is much easier when one of them is interocularly correlated ($\phi = 0$) than when one is interocularly anticorrelated ($\phi = 180$).

Experiment 2: ΔC_{diff} as a function of pedestal C_{diff}

In this experiment we measured ΔC_{diff} as a function of a pedestal C_{diff} for both the lower and upper ranges

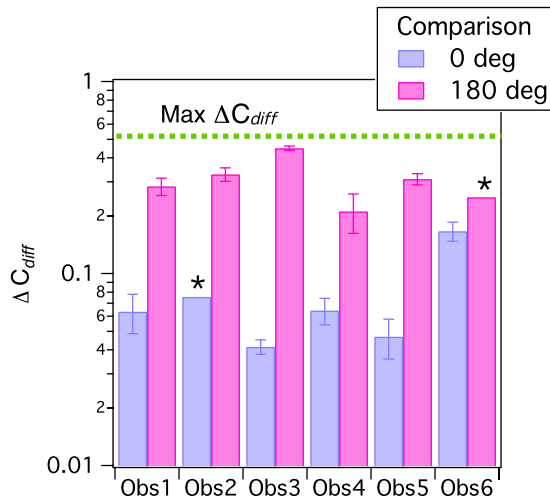Kingdom, Seulami, Jennings, & Georgeson

Figure 3. Results from Experiment 1. Blue bars show just-noticeable differences (ΔC_{diff}) for comparison stimuli that were correlated—i.e., had interocular phase difference $\phi = 0°$; magenta bars show just-noticeable differences for comparison stimuli that were anticorrelated—i.e., had an interocular phase difference $\phi = 180°$. Data for six observers. Green line shows the maximum possible threshold. Error bars in all graphs are bootstrap standard errors. Asterisks show cases where bootstrap errors could not be obtained, as also shown in the following data figures.

of C_{diff}. Figure 4 illustrates the conditions for the two parts of the experiment and the terms we will use for the graphical presentation of the data. As the figure shows, for the lower-range experiment the discriminand pairs comprised C_{diff} and $C_{\mathrm{diff}} + \Delta C_{\mathrm{diff}}$, which we designate respectively as the smaller and larger C_{diff} of a just-discriminable pair. For the upper-range data, the discriminand pairs are C_{diff} and $C_{\mathrm{diff}} - \Delta C_{\mathrm{diff}}$, which invites the opposite designation of, respectively, larger and smaller. The smaller-versus-larger designation enables the two sets of data to be directly compared in an intuitive manner.

Figure 5 presents on linear axes the larger versus smaller C_{diff} results for both lower- (blue) and upper-range (magenta) data. Thus, these graphs plot on the two axes the just-discriminable C_{diff} pairs across the full range of C_{diff}. Note that the orientations of the error bars differ for the two ranges. This reflects the fact that for the lower-range data, ΔC_{diff} varies along the ordinate, as it is part of the larger C_{diff}, but for the upper-range data it varies along the abscissa, as it is part of the smaller. For three observers (1, 2, 7) the just-noticeably larger C_{diff} rises smoothly with increasing values of the smaller C_{diff} until the physical limit (dashed green line) is reached, suggesting that these thresholds for the lower and upper ranges of C_{diff} lie on a single monotonic function. The contin-
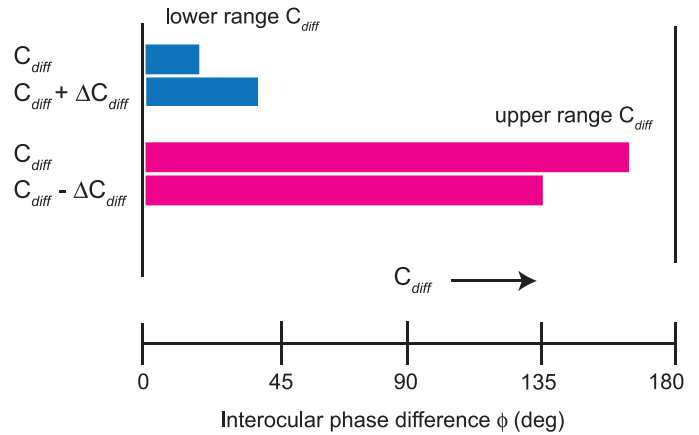

Figure 4. Protocol and measurement terms for Experiment 2. See text for details.

uous gray curves show fits of a B– model to be described later.

Figure 6 presents ΔC_{diff} as a function of the smaller C_{diff} on log-log axes. The continuous gray lines show fits of the same B– model as in Figure 5. The figure brings out more clearly the small dipper effect in the data at small levels of C_{diff}.

Discussion

We began with the hypothesis that JNDs for the lower and upper ranges of interocular difference should be similar because they would be mediated by binocular differencing (B–) and binocular summation (B+) channels, respectively. The results from both experiments, however, favor rejection of this idea. In the first experiment, observers found it much easier (by a factor of 4) to detect an interocular difference in C_{diff} when the comparison stimulus was interocularly correlated than when it was interocularly anticorrelated. This finding was further supported by a second experiment in which we measured JNDs across the full range of interocular difference.

If JNDs were mediated by B– channels in the lower range and B+ channels in the upper range, we would expect the pattern of JNDs to be mirror-symmetric around the midpoint of the range of interocular difference, yet the plots in Figure 6 show this not to be the case. In what follows, we show how a model based on just the B– channel is able to give a good account of the JND data.

B– channel model

Let us assume that the B– channel has a response function R that can be modeled similarly to the well-
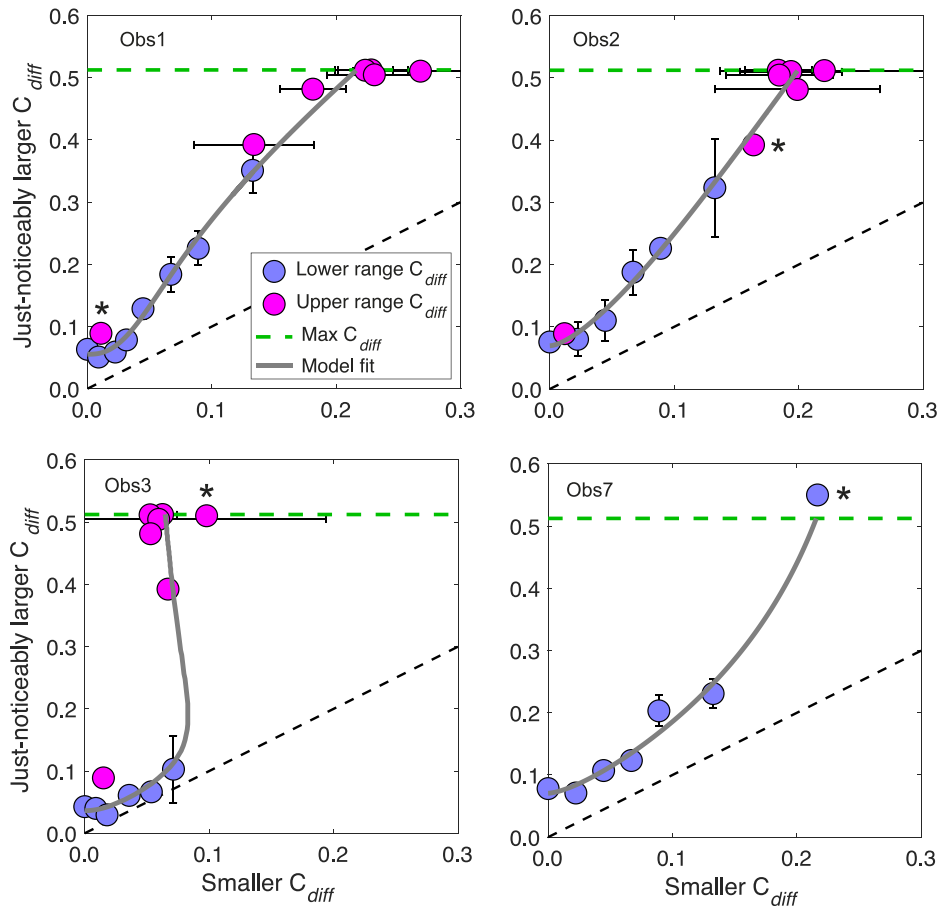
Figure 5. Just-noticeably larger C_{diff} as a function of smaller C_{diff}, for both the lower-range (blue) and upper-range (magenta) data, for four observers. Note that only the lower range was tested for Observer 7. Green dashed line shows the maximum C_{diff}. Diagonal black dashed line represents points of equal value on the two axes, and points representing just-noticeable differences must always lie above this line of equality. Note that the point in Observer 7's data that lies above the green line is there because the psychometric fitting procedure did not impose the maximum C_{diff} limit.

known contrast transduction model suggested originally by Legge and Foley (1980). In the terms of this study, the model is

$$R = \frac{C_{\text{diff}}^{p}}{\left(z + C_{\text{diff}}^{q}\right)}, \quad (3)$$

where p, q, and z determine the shape of the response function. With a suitable choice of parameters the function is able to capture the idea that R first accelerates and then decelerates as a function of C_{diff}, thus providing one possible explanation for the dipper function observed in our lower-range data. To apply the model, we assumed that the ΔC_{diff} at 75% correct elicits a constant threshold change in response ΔR, which we term k. This is equivalent to assuming that performance is limited by late additive noise with constant variance σ^2. The model's threshold ΔC_{diff} is then found for each pedestal C_{diff} by adjusting ΔC_{diff} until $\Delta R = k$. That is, for a given set of parameters p, q, z, k, we find ΔC_{diff} based on the following equation:

$$\Delta R = abs\{R([C_{\text{diff}} + task.\Delta C_{\text{diff}}], p, q, z) \\ - R(C_{\text{diff}}, p, q, z)\} = k, \quad (4)$$

where task $= 1$ for increments and -1 for decrements. Then, using the simplex algorithm (*fminsearch* in Matlab), we found the best-fitting parameter sets for each observer that minimized the sum of squared differences between model and observed thresholds. The best fit was taken over 50 repeated runs with jittered starting values, to avoid finding local minima in the error surface. The model fitting could minimize the squared error in either ΔC_{diff} or $\Delta\phi$. Because of the compressive relation between C_{diff} and ϕ (Figure 2), we chose to minimize errors in $\Delta\phi$. Table 1 gives the values of the fitted parameters and the coefficient of determination R^2 for each fit. Resulting model fits, re-expressed as C_{diff} or ΔC_{diff}, are the continuous gray lines in Figures 5 and 6.

The R^2 values indicate that this simple model, assuming a B− channel with a nonlinear transducer, gave a good fit to the data for all four observers. Figure
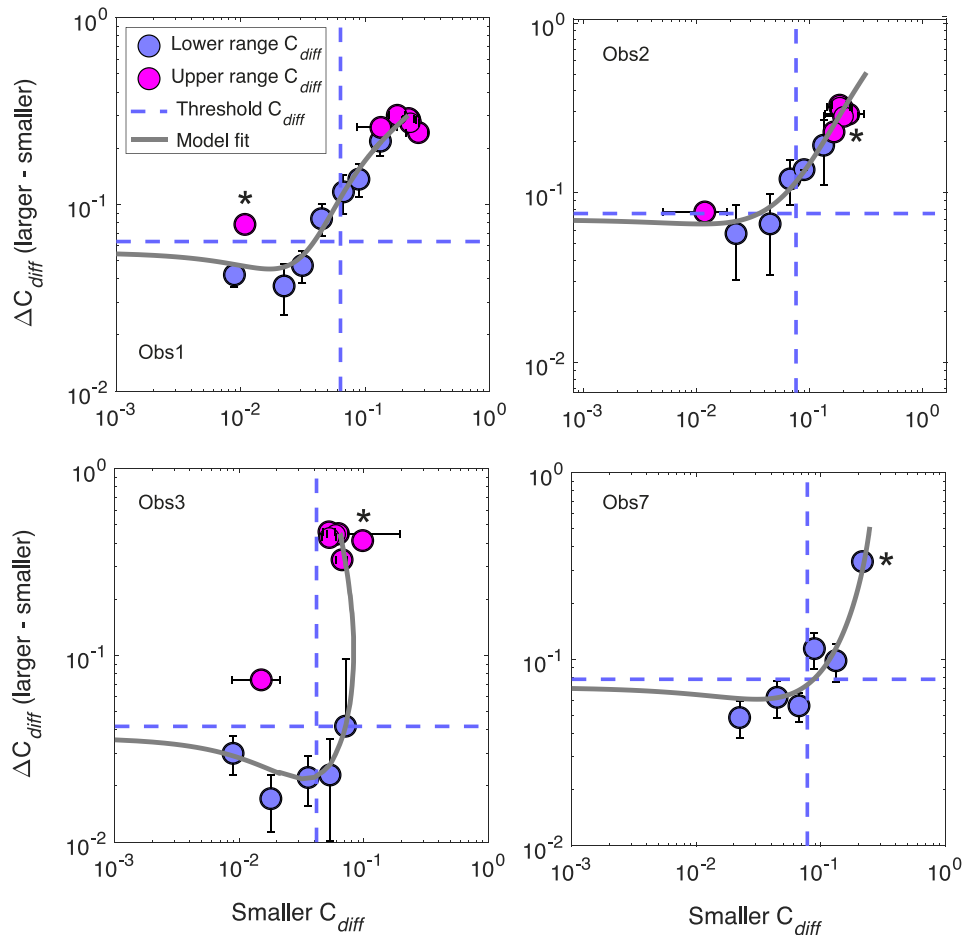
Figure 6. Same data as Figure 5, but with ΔC_{diff} plotted against the smaller C_{diff}, and on log-log axes. Dashed lines show threshold C_{diff}—i.e., the ΔC_{diff} value obtained when the lower-range comparison value of C_{diff} value was zero.

7 shows the estimated transducer functions for C_{diff} for the four observers. The functions are moderately compressive, and fairly similar for observers 1, 2, and 7. But observer 3 (FAAK) shows a more extreme, nonmonotonic transducer with saturation followed by decline for $C_{diff} > 0.2$, corresponding to the very steep rise in thresholds seen in Figures 5 and 6. In Appendix 1 we consider whether the nonmonotonic transducer for Observer 3 might be the result of overfitting (having more free parameters than are warranted by the data). For completeness, we report the results of an Akaike information criterion model-selection analysis, comparing three versions of the transducer model applied to the data from each observer. We conclude that the

transducer shapes shown in Figure 7 are not distorted by overfitting.

In this analysis, the saturating nonlinearity is modeled as occurring after the contrasts of the signals from the two eyes have been linearly differenced by the
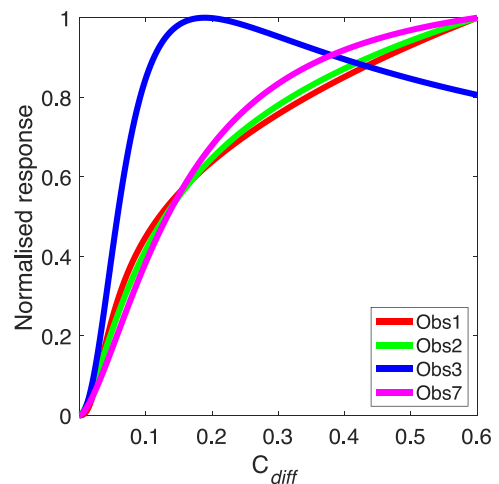

Figure 7. Estimated transducer shapes for the B− model for the four observers, normalized to their maximum values.

Observer	p	q	z	k	R^2
1	2.677	2.287	0.000565	0.228	0.991
2	1.608	1.365	0.0277	0.252	0.996
3	1.967	2.267	0.00349	0.365	0.997
7	1.498	1.536	0.063	0.236	0.990

Table 1. Parameter estimates for the model which gave the fits shown in Figures 5 and 6. See text for details.

B– channel. It is interesting to ask (as one referee did) whether our results could instead be explained by a nonlinearity applied to the monocular contrast signals prior to combination by a linear B– channel, as in Jennings and Kingdom's (2016) model of the B– channel, for example. Would an early nonlinearity seriously affect our reasoning about the functioning and primary role of the B– channel in the discrimination tasks studied here? In Appendix 2 we show that a nonlocal contrast nonlinearity applied prior to binocular differencing has no effect on the predictions of the B– channel for the experiments modeled here. We are therefore confident that while an early contrast nonlinearity almost certainly occurs prior to binocular differencing, it is insufficient to account for the results of the present study.

The properties of the B– channel revealed are relevant to early studies by Julesz and Tyler (1976) and Tyler and Julesz (1976, 1978; see also Julesz, 1986). Using dynamic random-dot correlograms, they found that across three observers the time needed to perceive a change from interocular correlation ($r = 1$) to uncorrelation ($r = 0$) was as brief as 5–10 ms, while the reverse direction, from interocular uncorrelation to correlation, required about five times as long (25–50 ms). Julesz and Tyler coined the term *neurontropy* to characterize this asymmetry, the idea being that in terms of entropy the switch from correlation to uncorrelation was one of order to disorder, while the reverse was one of disorder to order. Creating order may necessarily be a slower, more difficult process than reducing order to chaos. Julesz and Tyler proposed the interaction of two processes, fusion and rivalry, that operate in parallel and are akin to the B+ and B– channels, respectively. But that model did not directly account for the striking difficulty of detecting an increase compared to a decrease of correlation. Tyler and Julesz (1978) suggested that this difference "must represent some kind of *adaptation* to the current state" of the visual noise (p. 104). This is consistent with our previous study showing the B– channel to be adaptable (Kingdom et al., 2018). In the Tyler/Julesz experiments the transitions would be detected under different levels of B– channel adaptation. Transition from correlated ($r = 1$) to uncorrelated ($r = 0$) would be detected by a B– channel in an unadapted and hence maximally sensitive state, while the opposite transition ($r = 0$ to $r = 1$) would be detected by the same channel in an adapted and hence less sensitive state.

The B– channel and efficient coding

In the Introduction we mentioned a number of studies providing support for the existence of a binocular differencing channel. Some of these studies were motivated by a recent theory of binocular vision advanced by Li and Atick (1994) and Zhaoping (2014; Li and Zhaoping are the same person: Zhaoping Li) suggesting that early in vision, the retinal images of the two eyes are processed by two binocular channels: B+ which sums their signals, and B– which differences them. Crucially, the two channels are subject to separate gain controls. The idea is that the B+ and B– channels constitute an efficient code for representing binocular information, since they serve to decorrelate the highly correlated left- and right-eye signals. As mentioned in the Introduction, there is evidence for involvement of the B– (and B+) channels in a variety of visual tasks. From our finding that JNDs in C_{diff} appear to be mediated by the same mechanism across the full range of C_{diff}, we suggest that this mechanism is the B– channel.

Why not the B+ channel?

Stevenson et al. (1992) used dynamic random-dot stereograms to quantify the ability to detect small amounts of interocular correlation, specifically departures from zero correlation. In the terms of the present study this would translate to measuring decrement JNDs in C_{diff} with the comparison stimuli at 90° phase difference. Stevenson et al. found that thresholds were elevated by adapting to perfectly correlated images, complementary to our later finding that adapting to uncorrelated as well as anticorrelated images raised thresholds for detecting departures from correlation (Kingdom et al., 2018). The impairment of correlation detection shown by Stevenson et al. was disparity specific: The largest effect was obtained when the test correlation (embedded in uncorrelated noise) had the same disparity as the adapter. They interpreted these results as due to adaptation of disparity-tuned neurons within the stereovision system. This seems very likely, but also suggests that the B+ channel (in addition to the B– channel) might in principle be involved.

The results of the present study, however, suggest otherwise. It is worth reiterating that our use of horizontally oriented grating patterns prevented the use of stereo-depth cues to determine the stimulus with the larger C_{diff}. Informal observations by FAAK suggest that, unlike these gratings, stereo cues are quite pronounced in orientationally broadband stimuli with differing amounts of C_{diff}. So it is possible that the apparent lack of B+ involvement in the present study in comparison to previous related studies (e.g., Stevenson et al. 1992) may be due to the particular stimuli we used.

This leaves open the question why a horizontally oriented B+ channel is *not* involved. Recently,

Georgeson et al. (2016) put forward a model of binocular combination to account for the appearance of dichoptic mixtures of luminance contrasts and discrimination-threshold measures obtained in dichoptic masking experiments. They suggested that three channels were involved, two monocular (call these L and R) alongside the binocular summing B+ channel. The critical model computation was that the task-relevant visual response was given by the channel with the largest of the three outputs—that is, MAX(L,R,B+). This MAX operation can be envisaged as a form of competition or winner-take-all rivalry between all three signals. Such an operation is *not* needed to explain simple contrast increment discriminations, either monocular or binocular, but was strongly implicated in tasks where the contrast of a binocular pedestal was incremented in one eye and decremented in the other (see Georgeson et al., 2016, figure 9). It is also consistent with the near-winner-take-all behavior of binocular contrast matching. Ding, Klein, and Levi (2013) and Ding and Levi (2016, 2017) developed a detailed alternative account of binocular combination based solely on a more complex B+ channel, involving several interacting contrast-gain controls. Because it generates, by different means, a near-winner-take-all binocular response surface that fits contrast-matching behavior, the Ding model is likely to be able to predict correctly the critical contrast-decrement discriminations mentioned earlier (see Georgeson et al., figure 6A). But since it is focused entirely on the B+ summation mechanism, it will probably need additional mechanisms for interocular difference detection.

In the present study, the MAX operation provides a plausible explanation for the apparent absence of the B+ response: In our anticorrelated baseline conditions, the B+ channel would be overwhelmed by signals from the two monocular channels. To see why, recall that the binocular contrast C_{sum} falls to zero with increasing phase disparity (Figure 1), while the monocular contrasts C_L, C_R are invariant with disparity. Thus it seems likely that in the MAX operator, the B+ signal must be silenced by the L, R signals at *large* disparities after some critical disparity is reached. At *smaller* disparities, C_{sum} is larger, and the B+ signal may win the L,R,B+ competition, depending on how much binocular summation B+ exhibits. But that larger B+ signal is likely to be highly compressed (Figure 1), and so its ability to signal changes may be much lower than the B− signal. Thus the B+ signal may fail to support discrimination for different reasons at small and large disparities. To illustrate and quantify this argument, we developed a simple multichannel model that includes B−, B+, and MAX signals, as follows.

Discrimination by the B− channel despite high noise

To compare the possible roles of B− and B+, it is useful to work with a common input variable—the component phase disparity of our multicomponent images. Figure 8A and 8B use elementary signal-detection theory to summarize how the proposed nonlinear response of the B− channel accounts for the discrimination of changes in interocular difference, expressed here as phase disparity. Thin curves show the two binocular contrast signals that might be used in these tasks: C_{diff} (solid) and C_{sum} (dashed), as in Figure 2. The thick gray curve shows the fitted response of the B− channel (for Observer 1) after nonlinear transduction of C_{diff} (Equation 3). Responses to the pedestal levels of disparity are marked on this model response curve as white circles. An increase in phase disparity (hence an increase in C_{diff}) raises the B− response, and discrimination threshold is reached when the rise ΔR equals a constant k (Equation 4) corresponding to 75% correct, $d' = 0.95$. Because d' is nearly 1, and $d' = \Delta R/\sigma$, it follows that k almost equals the internal noise level σ ($k = 0.95\sigma$).

For our observers, k was around 0.25 (Table 1), but since the full range of model responses was only about 0.8 (Figure 8A), we can see that the task is noisy: Its full range spans only about three standard deviations of the noise σ. Colored circles represent the observed test threshold values (expressed as phase disparity, $\phi_{\mathrm{ped}} + \Delta\phi$) tied to the corresponding pedestal levels (in three selected cases) by black line segments. Figure 8B is similar to Figure 8A, but for discrimination of decreases in phase disparity (or C_{diff}). Importantly, in both Figure 8A and 8B, the horizontal positions of these threshold points are empirical, not model dependent, while their vertical displacement from the pedestal points is the constant k. The fact that all these threshold points fall on or very close to the same model response curve tells us that the fitted transducer for B− accounts very well for the discrimination of both increases and decreases in interocular phase difference.

Another sign of the low signal-to-noise ratio in this task was the high value of Weber fractions: For increments in C_{diff}, the values of $\Delta C/C$ for observers 1, 2, and 7 were about 1.0, 1.5, and 1.5, meaning that the just-detectable increment was as large as or larger than the pedestal itself. These Weber fractions are about five to 10 times higher than for grating-contrast increment detection, where $\Delta C/C$, monocularly or binocularly, is typically about 0.1 to 0.2 (see, e.g., Georgeson et al., 2016, figure 4A, 4B, and 4C). This greater Weber fraction probably reflects a much higher noise level in the C_{diff} task, because the transducers for C_{diff} and for grating contrast were broadly similar in shape (cf. Figure 9). How much of the excess noise in the C_{diff} task might be due to the noisy nature of our compound gratings (the random phase relation between components) is not yet known.
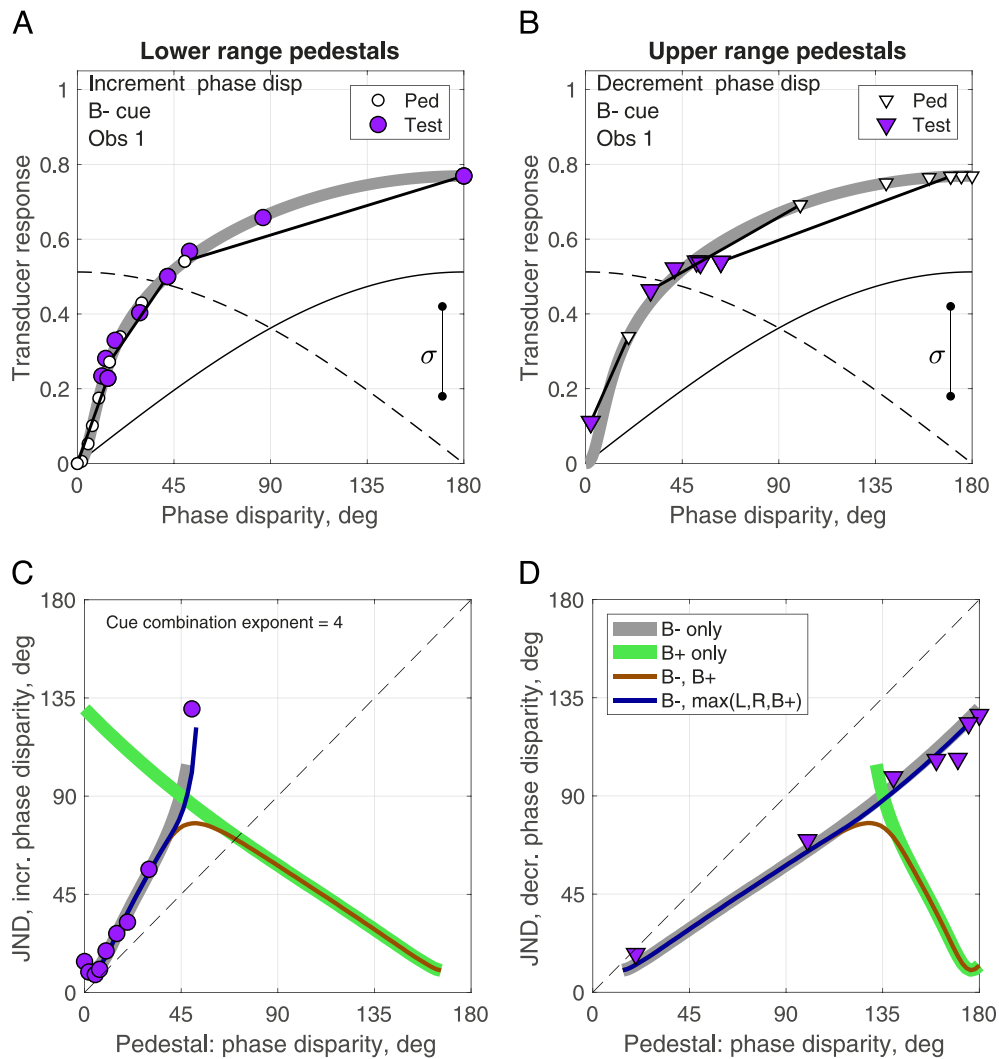
Kingdom, Seulami, Jennings, & Georgeson

Figure 8. How the B− channel accounts for discrimination of increases and decreases in interocular difference, and possible reasons why the B+ channel does not contribute. (A) Thin black curves are C_{diff} (solid) and C_{sum} (dashed) as a function of component phase disparity, as in Figure 2. Thick gray curve is the fitted response of the B− channel (for Observer 1) after nonlinear transduction of C_{diff} (Equation 3). Model responses to pedestal disparity (ϕ_{ped}) are marked on this curve as white circles. Colored symbols represent experimental test threshold values (expressed as phase disparity, $\phi_{\text{ped}} + \Delta\phi$); in three examples these are tied to their corresponding pedestal levels by black line segments. Internal noise level σ derived from the model fit is marked by a vertical bar. (B) As in (A), but for discrimination of decreases in phase disparity. (C, D) Symbols are just-noticeable difference for (C) an increase in phase disparity or (D) a decrease, as a function of ϕ_{ped}. Thick gray curves show the good fit of the B− channel alone, in both cases. Thick green curves show that B+ predictions bore no resemblance to the data. In a model allowing both cues to be used, the more sensitive of the two cues (B−, B+; thin brown curve) worked well only where B− was the better cue; the B+ contribution elsewhere was far too strong. However, when B+ was in competition with monocular signals L, R, its predicted contribution to the task was almost nil (thin blue curve; see Figure 9, left); hence, B− was the only effective cue.

Is the B+ channel vetoed by monocular signals?

We first attempted to fit the data assuming only a B+ channel, again using Equations 3 and 4, but inserting C_{sum} in place of C_{diff}. The simplex fitting algorithm did not converge on any set of transducer parameters or noise level that could emulate the data, even approximately. Thus, as expected, it seems likely that B+ was not used. To get further insight into why not, we made the simplifying assumption that, for a given observer,

the transducers T for L, R, and B+ were the same as for B−. Thus the four channel responses were $R_L = T(C_L)$, $R_R = T(C_R)$, $R_{\text{sum}} = T(C_{\text{sum}})$, and $R_{\text{diff}} = T(C_{\text{diff}})$. The MAX operator response was then defined by a Minkowski sum with a high exponent:

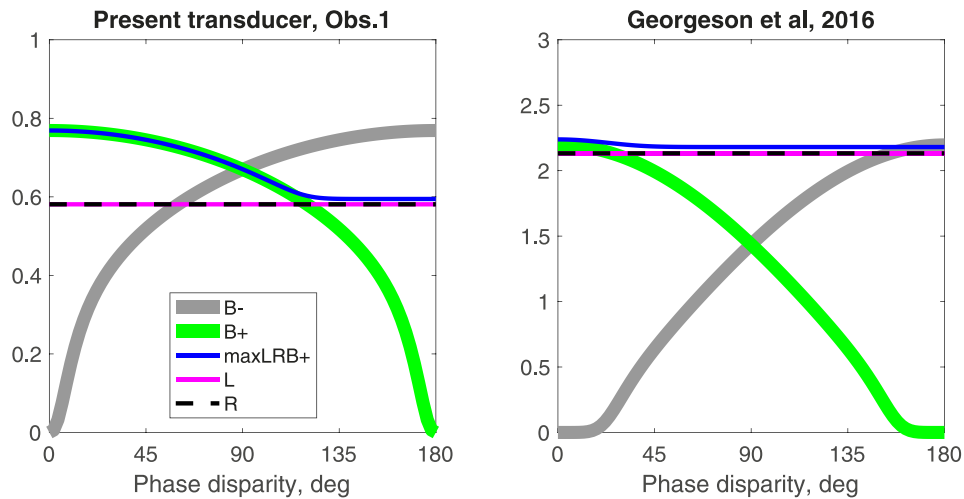$$R_{\text{max}} = \left(\sum_{i=L,R,\text{sum}} R_i^n \right)^{\frac{1}{n}},$$

Figure 9. Left: How the MAX operator output (thin blue curve) enables monocular responses (red, black) to occlude the binocular B+ response (green) at larger disparities. At smaller disparities, any changes in B+ or MAX response with disparity would be below threshold since the changes are smaller than the noise level (about 0.25). Right: A model based closely on the two-stage model equations and parameter values of Georgeson et al. (2016) shows even greater exclusion of the B+ response by the MAX operation, because (by design, and by model fitting to discrimination data) the monocular responses were almost as large as the largest B+ response. This comparison of models adds some generality to the idea that the binocular summing channel would not contribute to performance in the discrimination of interocular differences.

where $n = 30$ (Georgeson et al., 2016). We then computed d' for each channel alone, or in combination, as a function of phase disparity, where for the ith channel $d'_i = \Delta R_i / \sigma$, and ΔR_i is defined as in Equation 4, with the appropriate change of contrast variable. To combine d' values across channels i, j, we again used a Minkowski sum:

$$d'_{OBS} = \left(d'^m_i + d'^m_j \right)^{\frac{1}{m}},$$

where $m = 4$. A value of $m = 2$ represents optimal combination for statistically independent cues (the ideal observer; Green & Swets, 1966), but this is unachievable in practice because it requires the observer to have perfect knowledge of the signal means and their detectabilities d' on each trial in order to weight the cues optimally. A weaker form of summation ($m = 4$) seems appropriate, and is not crucial to our argument. The resulting d' tends to track the higher of the two d' values but shows some summation when the two d' values are similar. In this way we computed the expected increment thresholds $\Delta\phi$ for single cues (B− or B+) and for pairs of cues—(B−, B+) and (B−, R_{max})—as shown in Figure 8C and 8D. Thick curves represent the single-cue predictions, as indicated.

Because of the symmetry between C_{sum} and C_{diff} (thin curves in Figure 8A and 8B), it follows that the threshold curve for B+ with disparity decrements (green curve, Figure 8D) is the mirror image of the B− curve for disparity increments (gray curve, Figure 8C). There is an analogous symmetry between B+ with increments

(Figure 8C) and B− with decrements (Figure 8D). But only the B− curves fit the data. With the cue combination (B−, B+), predicted thresholds (thin brown curves) track the better cue across the whole range of pedestal disparities. But the observed thresholds did not do this for any observer. Finally, when the (B−, R_{max}) cues were combined (thin blue curves), predicted thresholds reverted to being very close to those for B− alone, and close to the data. B+ failed to deliver useful information because the response R_{max} varied so little with phase disparity (see Figure 9, left, thin blue curve) in relation to the noise level.

We conclude from this analysis that the B+ signal probably plays no part in these discriminations because it is occluded by monocular signals at large disparities and has insufficient discriminative capacity at small disparities. This conclusion must be tentative because B+ is effectively silenced, so we have no direct evidence about the form of the B+ transducer from these experiments. Nevertheless, the same conclusion can be drawn from applying the model of Georgeson et al. (2016). Here (Figure 9, right) the L, R, B+, and MAX response curves are based directly on their model and parameters which they had fitted to binocular contrast-discrimination data. The inability of B+ to pass information about interocular difference through the MAX operator (thin blue curve) is even more evident. Finally, because B− responses *are* an effective cue for interocular difference detection, it follows that they do *not* pass through a MAX operator in competition with monocular responses. This is

broadly in agreement with the previous proposal that a luster signal operates in parallel with the MAX signal (Georgeson et al., 2016).

Conclusions

We have provided compelling evidence that the detection of interocular differences in grating phase, and hence local contrast, is mediated exclusively by a B– channel, in spite of the fact that for a range of conditions the B+ channel on a priori grounds would be expected to mediate detection. We suggest that this lack of B+ contribution occurs because the B+ channel output is vetoed when signals from the monocular channels are stronger.

Keywords: interocular difference, interocular contrast, binocular differencing channel, binocular summing channel

Acknowledgments

Corresponding author: Frederick A. A. Kingdom.
Email: fred.kingdom@mcgill.ca.
Address: McGill Vision Research, Department of Ophthalmology, Montréal General Hospital, Montréal, Canada.

References

Baker, D. H., Lygo, F. A., Meese, T. S., & Georgeson, M. A. (2018). Binocular summation revisited: Beyond $\sqrt{2}$. *Psychological Bulletin, 144*(11), 1186–1199.

Baker, D. H., Wallis, S. A., Georgeson, M. A., & Meese, T. S. (2012). Nonlinearities in the binocular combination of luminance and contrast. *Vision Research, 56*, 1–9.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference* (2nd ed.). New York: Springer.

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

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research, 31*, 2195–2207.

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, https://doi.org/10.1167/13.2.13. [PubMed] [Article]

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

Ding, J., & Levi, D. M. (2017). Binocular combination of luminance profiles. *Journal of Vision, 17*(13):4, 1–32, https://doi.org/10.1167/17.13.4. [PubMed] [Article]

Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proceedings of the National Academy of Science USA, 103*, 1141–1146.

Formankiewicz, M. A., & Mollon, J. D. (2009). The psychophysics of detecting binocular discrepancies of luminance. *Vision Research, 49*(15), 1929–1938.

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 Research, 129*, 98–118.

Goncalves, N. R., & Welchman, A. E. (2017). "What not" detectors help the brain see in depth. *Current Biology, 27*, 1403–1412.

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

Hovis, J. K. (1989). Review of dichoptic color mixing. *Optometry and Vision Science, 66*, 181–190.

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, https://doi.org/10.1167/16.3.23. [PubMed] [Article]

Julesz, B. (1986). Stereoscopic vision. *Vision Research, 26*, 1601–1612.

Julesz, B., & Tyler, C. W. (1976) Neurontropy, an entropy-like measure of neural correlation in binocular fusion and rivalry. *Biological Cybernetics, 23*, 25–32.

Katyal, S., Vergeer, M., He, S., He, B., & Engel, S. A. (2018). Conflict-sensitive neurons gate interocular suppression in human visual cortex. *Scientific Reports, 8*, 1239.

Kingdom, F. A. A. (2012). Binocular vision: The eyes add and subtract. *Current Biology, 22*, R22–R24.

Kingdom, F. A. A., Jennings, B. J., & Georgeson, M. A. (2018). Adaptation to interocular difference. *Journal of Vision*, *18*(5):9, 1–11, https://doi.org/10.1167/18.5.9. [PubMed] [Article]

Kingdom, F. A. A., & Libenson, L. (2015). Dichoptic saturation mixture: Binocular luminance contrast promotes perceptual averaging. *Journal of Vision*, *15*(5):2, 1–15, https://doi.org/10.1167/15.5.2. [PubMed] [Article]

Kingdom, F. A. A., Yared, K.-C., Hibbard, P., & May, K. (in press). Stereoscopic depth adaptation from binocularly correlated versus anti-correlated noise: Test of an efficient coding theory of stereopsis. *Vision Research*.

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

Li, Z., & Atick, J. J. (1994). Efficient stereo coding in the multiscale representation. *Network: Computation in Neural Systems*, *5*, 157–174.

Malkoc, G., & Kingdom, F. A. A. (2012). Dichoptic difference thresholds for chromatic stimuli. *Vision Research*, *62*, 75–83.

May, K. A., & Zhaoping, L. (2016). Efficient coding theory predicts a tilt aftereffect from viewing untilted patterns. *Current Biology*, *26*, 1571–1576.

May, K., Zhaoping, L., & Hibbard, P. (2012). Perceived direction of motion determined by adaptation to static binocular images. *Current Biology*, *22*, 28–32.

Prins, N., & Kingdom, F. A. A. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. *Frontiers in Psychology*, *9*: 1250.

Reynaud, A., & Hess, R. F. (2018). Interocular correlation sensitivity and its relationship with stereopsis. *Journal of Vision*, *18*(1):11, 1–11, https://doi.org/10.1167/18.1.11. [PubMed] [Article]

Said, C. P., & Heeger, D. J. A. (2013). Model of binocular rivalry and cross-orientation suppression. *PLoS Computational Biology*, *9*, e100299.

Stevenson, S. B., Cormack, L. K., Schor, C. M., & Tyler, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vision Research*, *32*, 1685–1694.

Symonds, M. R. E., & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, *65*(1), 13–21.

Tyler, C. W., & Julesz, B. (1976). The neural transfer characteristic (neurontropy) for binocular stochastic stimulation. *Biological Cybernetics*, *23*(1), 33–37.

Tyler, C. W., & Julesz, B. (1978). Binocular cross-correlation in time and space. *Vision Research*, *18*, 101–105.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*(1), 192–196.

Yoonessi, A., & Kingdom, F. A. A. (2009). Dichoptic difference thresholds for uniform color changes applied to natural scenes. *Journal of Vision*, *9*(2):3, 1–12, https://doi.org/10.1167/9.2.3. [PubMed] [Article]

Zhaoping, L. (2014). *Understanding vision: Theory, models, and data*. Oxford, UK: Oxford University Press.

Appendix 1: Are the transducer shapes distorted by over-fitting the data?

In the main text (Figures 5 through 7) we fitted a standard four-parameter model (Equation 3, with parameters p, q, z plus the noise parameter k) widely used in the context of luminance contrast discrimination, and here applied to C_{diff}. Let's call this Model 1. For three observers this gave rise to a smoothly saturating transducer function, but for Observer 3 the fitted transducer was, surprisingly, nonmonotonic (Figure 7). We therefore aimed to determine whether a more constrained three-parameter model might also fit the data well, without an unusual transducer shape. We considered two reduced versions of Equation 3: In Model 2, q was fixed ($q = 2$) while p was free to vary in the model fitting; in Model 3, p and q were yoked (i.e., varied together, $p = q$). A powerful general procedure for comparing the goodness of different models, especially when they are not nested, is based on the AIC (Akaike information criterion; e.g., Burnham & Anderson, 2002; Wagenmakers & Farrell, 2004; Symonds & Moussalli, 2011). The AICc (AIC corrected for small samples) takes into account both the goodness of fit (deviance or squared error) and the model complexity (number of parameters), and returns Akaike weights that can be interpreted as the probability that a given model is the best of those considered. The outcome of the AIC analysis is shown in Table A1. We can see that for each observer, one of the three-parameter models emerged as best (i.e., more likely to be closest to an unknown true model). For two observers it was Model 2, while for the other two it was Model 3. Clearly it is undesirable to select different
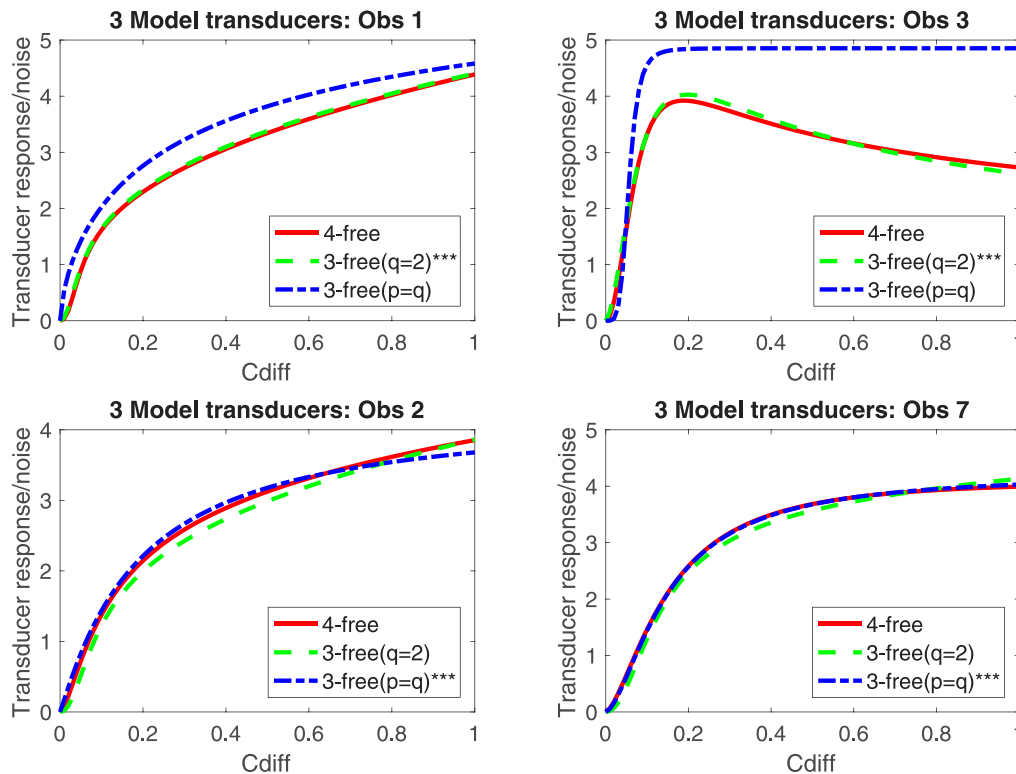
Figure A1. Each panel shows the form of the transducer produced by three different but related models fitted to the data for a given observer. The models were variants of the same transducer equation (Equation 3), and the ordinate plots R/k, the transducer response R scaled by the model noise parameter k. For each observer, the best model, as returned by the Akaike information criterion analysis, is marked by *** in the legend. The best model was always one or other of the three-parameter models, but we note that the four-parameter model was in each case very close to the best model.

models for different observers. But we note that in all four cases (Figure A1) the original model (Model 1) gave transducer functions (red curves) that were very close to those of the best model. It seems reasonable to conclude that Model 1 is a suitable model for all four observers, and that the extra flexibility gained from the

fourth parameter did not lead to a distortion in the shape of the transducer function to any serious extent, compared with alternative models that were less flexible.

Appendix 2: Are there monocular nonlinearities before interocular difference detection?

In our model the B– channel was assumed first to respond to the linear difference of the spatial-contrast waveforms in the two eyes, followed by a nonlinear, compressive transformation of the contrast of this combined waveform. Here we ask what might happen if, as seems likely, the monocular responses were a nonlinear function of the two contrasts C_L, C_R before the differencing operation. We show here that for our experiments, and for a broad class of possible non-linearities, such models would give exactly the same predictions as the model with a linear front end that we described, but further experiments could shed new light on this question.

Observer	Model #	AICc	ΔAIC	Akaike weight	Best model
1	1	59.2606	3.608	0.139	
	2***	55.6522	0	0.844	3-free ($q = 2$)
	3	63.4882	7.836	0.017	
3	1	48.3753	4.179	0.091	
	2***	44.1962	0	0.736	3-free ($q = 2$)
	3	47.0860	2.89	0.173	
2	1	40.4936	2.794	0.129	
	2	38.5200	0.82	0.347	
	3***	37.6999	0	0.523	3-free ($p == q$)
7	1	47.0704	13.945	0.001	
	2	34.7354	1.61	0.309	
	3***	33.125	0	0.691	3-free ($p == q$)

Table A1. Akaike information criterion (AIC) comparison of three models. The best model for each observer is denoted by ***. *Notes*: AICc = AIC corrected for small samples.

Let the left- and right- eye inputs I_L, I_R be represented by

$$I_L = C_L f_L(x, \phi), \quad I_R = C_R f_R(x, \phi), \quad (A1)$$

where C_L, C_R are the RMS contrasts and f_L, f_R are the disparate spatial-contrast waveforms used in the experiments, with amplitude scaled to give each a standard deviation of 1, and component phase disparity ϕ. We then broadly follow the approach of Ding and colleagues (Ding & Sperling, 2006; Ding et al., 2013; Ding & Levi, 2017) and Jennings and Kingdom (2016) in supposing that multiplicative weights w_L, w_R alter the effective contrasts of these signals but do not alter the waveforms f_L, f_R before linear combination across the eyes. Thus,

$$r_L = w_L C_L f_L(x, \phi), \quad r_R = w_R C_R f_R(x, \phi), \quad (A2)$$

and their combination $r_{B-} = r_R - r_L$. The final response of the B– channel would then be

$$R_{B-} = T(std\{r_{B-}\}), \quad (A3)$$

where T is a nonlinear transducer function of the kind discussed in the main text and *std* returns a single number: the standard deviation of values over space x, or some other aggregate measure of response strength over space.

In general the weights will be a function W of C_L, C_R and other factors such as those related to luminance level (Ding & Levi, 2017), but for our experiment, where the RMS contrasts were always equal ($C_L = C_R$), we can strongly expect the weights to be equal. For example, a simple form of ocular weighting would be

$$w_L = \frac{C_L^\gamma}{C_L^\gamma + C_R^\gamma}, \quad w_R = \frac{C_R^\gamma}{C_R^\gamma + C_L^\gamma}, \quad (A4)$$

where γ is a constant exponent. Here $w_L = w_R = 0.5$. But notice how this equality depends on two things: the equality of contrasts and the left/right symmetry of the weight equations. Such symmetry seems very likely for normal observers with balanced ocular properties. Thus if, in general, $w_L = W(C_L, C_R, \alpha, \beta, \gamma)$, where $\alpha, \beta,$

γ are constants, then it follows from symmetry that $w_R = W(C_R, C_L, \alpha, \beta, \gamma)$. Hence, with equality of contrasts (and equality of other factors such as luminance level and spatial power spectrum), the weights must be equal, whatever the form of the weight equation W and no matter how complex it may become (e.g., Ding & Levi, 2017, model 5).

Given this equality of the weights and contrasts in our experiments, it follows from Equation A2 that

$$r_{B-} = k[C_L f_L(x, \phi) - C_R f_R(x, \phi)], \quad (A5)$$

where $k = w_L = w_R$ and will be the same for all conditions of the experiments, provided W does not vary with phase disparity. Within these broad constraints, we can see that even with arbitrarily complex monocular weight functions W, which may incorporate contrast nonlinearities and interocular suppression, the binocular difference r_{B-} is equal to the linear difference between the input stimuli, up to a constant scaling factor k. More particularly, Equation A5 implies that the standard deviation of r_{B-} is directly proportional to C_{DIFF}, and from Equation A3 we get the B– response strength $T(std\{r_{B-}\}) = T(k.C_{DIFF})$. In short, for these experiments, the predictions of a linear differencing front end would not be altered by such nonlinear weighting, provided the output transducer function T were (trivially) rescaled to allow for the factor k change in input amplitude. This initially seemed surprising, so to confirm our reasoning we ran model fits with the linear front end as usual or with weighting schemes such as Equation A4. The fitted curves and goodness of fit to the data were indistinguishable.

This conclusion does *not* demonstrate that the front-end differencing is linear, only that we could not determine what monocular nonlinearities, if any, were present. This uncertainty arises because the left and right contrasts were constant and equal. Future experiments in which left and right contrasts are systematically varied would shed new light on the monocular weights that precede interocular difference detection.