A simple model of binocular luster

Gunnar Wendt Institut für Psychologie, Universität Kiel, Kiel, Germany

Franz Faul Institut für Psychologie, Universität Kiel, Kiel, Germany

The dichoptic combination of simple center–surround stimuli showing a contrast difference between eyes can trigger a lustrous impression in the fused percept, particularly when the contrast polarities in the two input images are of opposite sign. Recent developments suggest that the phenomenon of binocular luster results from a neural conflict between ON and OFF visual pathways at an early binocular level. Support for this idea was found in a previous study in which the empirical luster judgments strongly correlated with the predictions of an interocular conflict model which was based on such ON–OFF pairings. However, our original model could not account for the fact that weaker lustrous sensations can also be evoked by stimuli showing contrast polarities of same sign between eyes. In the present study we present an improved model that also includes ON–ON and OFF–OFF pairings. The predictive power of this model was tested in a series of four experiments, using a total of about 500 different center–ring–surround configurations as test stimuli. We found that, overall, our modified version accounts for more than 80% of the variance in the empirical luster judgments and that the former problems could be largely resolved. Our results further suggest a nonlinear transducer function for the binocular conflict signals.

Introduction

The phenomenon of binocular luster, which can be produced with simple dichoptic center–surround configurations showing a luminance contrast difference between corresponding center patches, has been studied with varying intensity during the past 170 years (Wendt & Faul, 2022; see also Mausfeld, Wendt, & Golz, 2014). Several theories have been proposed as an explanation for the occurrence of a lustrous impression in these stimuli, most notably the Oppel–Helmholtz theory of unconscious inference (Oppel, 1854; Oppel, 1857; von Helmholtz, 1856; von Helmholtz, 1867). According to this approach, a lustrous sensation is generated by the visual system in these cases because the presence of interocular luminance differences at corresponding retinal locations is interpreted as being caused by an object with specular and therefore directionally selective reflection.

In recent years, however, an alternative theory of binocular luster, which had originally been proposed by Brewster (1861), received strong support from several psychophysical studies. In particular, Anstis (2000) could show that the lustrous response is considerably stronger when the contrast polarities between corresponding center–surround stimuli are reversed than when they are of the same sign. That is, the interocular combination of a luminance increment with a luminance decrement (inc–dec pairing; see Figure 1, left) produces a much stronger impression of luster in the fused percept than stimuli in which different increments (inc–inc pairing; see Figure 1 middle) or different decrements (dec–dec pairing) are dichoptically combined (Chua, Zhang, Hammad, Zhao, Goyal, & Singh, 2015; Venkataramanan, Gawde, Hathibelagal, & Bharadwaj, 2021; Wendt & Faul, 2019; Wolfe & Franzel, 1988). Anstis (2000) assumed that the lustrous effect results from the inability of the visual system to binocularly combine the two discrepant monocular contrast signals: At the retina, incremental and decremental luminance patterns are encoded by two different types of ganglion cells, which are referred to as ON- and OFF-center cells, respectively (see also the model description in the following section). A dichpotic inc–dec stimulus would trigger these two cell types at corresponding retinal positions whose signals, however, cannot be integrated at a binocular level and therefore remain in a state of conflict, causing a sensation of luster (Burr, Ross, & Morrone, 1986; Sachsenweger, 1960).

Further support for this *conflict* approach based on ON- and OFF-center cell interactions comes from a recent study by Wendt and Faul (2020). In their experiment, they measured the magnitude of the lustrous sensation elicited by center–ring–surround stimuli in which the luminances of the ring and the surround, as well as the width of the ring, were varied while the interocular luminance difference between corresponding center patches was kept fixed. Following the idea from Anstis (2000), Wendt and Faul (2020) developed a simple *interocular conflict* model based on

Citation: Wendt, G., & Faul, F. (2022). A simple model of binocular luster. *Journal of Vision*, *22*(10):6, 1–26, https://doi.org/10.1167/jov.22.10.6.

https://doi.org/10.1167/jov.22.10.6 Received May 18, 2022; published September 8, 2022 ISSN 1534-7362

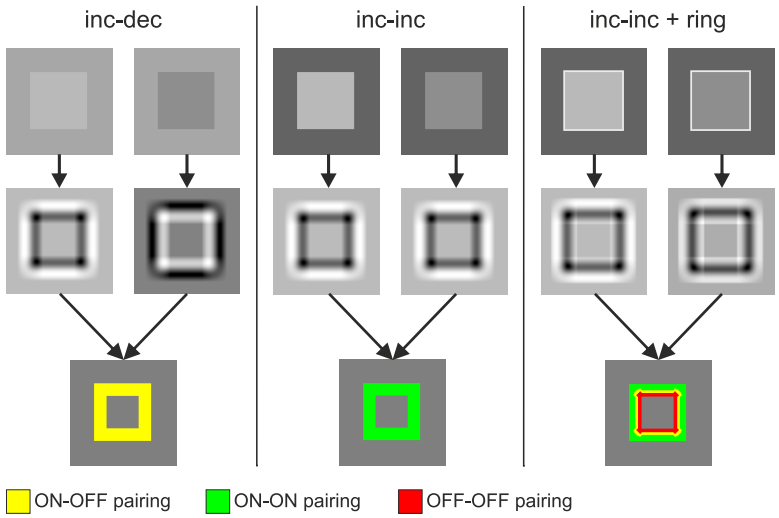

Figure 1. Three different dichoptic stimulus pairs that produce different sets of interocular contrast signal combinations and, as a consequence, different degrees of luster (columns). In order to identify the different types of interocular contrast pairings, the two half-images of a stimulus pair (top row) are first convolved with an LoG filter (second row; in the examples, a filter radius of 14 pixels was used). A stimulated ON-center cell at a given pixel position is represented by a negative filter value (dark pixels in the filtered images), and a stimulated OFF-center cell accordingly by a positive filter value (light pixels). The bottom row shows the occurrence of the different types of interocular contrast signal pairings within the target area. The inc–dec stimulus (left), which is perceived as strongly lustrous, exclusively produces ON–OFF pairings (yellow area). The inc–inc stimulus (middle) will only trigger ON-center cells at corresponding retinal positions; therefore, only ON–ON pairings (green areas) occur at the binocular stage, giving rise to a rather weak impression of luster. The ring stimulus (right) produces a much more complex pattern of contrast signal combinations, which includes ON–OFF, ON–ON, and OFF–OFF (red area) pairings. This stimulus evokes a much stronger lustrous impression than the inc–inc stimulus. Note that, in the central part of the target areas, neither ON- nor OFF-center cells are stimulated, because the receptive fields only cover the uniform center patch area of the stimulus but no luminance edges. Nevertheless, the entire target area is perceived as lustrous. This is explained by the assumption that the lustrous quality originates at the edge between target area and neighboring luminance elements and then spreads into the remaining parts of the target area via filling-in processes (Wendt & Faul, 2020; Zöller, 1998).

such ON–OFF pairings (Figure 1, left) and found that it well predicts the empirical luster judgments made in the experiment.

However, despite a strong overall correlation between luster judgments and model predictions, there were some stimulus conditions for which the predictions seemed to be systematically wrong. In particular, as was also found in a number of other studies (Anstis, 2000; Formankiewicz & Mollon, 2009; Sheedy & Stocker, 1984; Wendt & Faul, 2019; Zhang, 2015), some weaker lustrous impressions were evoked by simple inc–inc and dec–dec stimulus combinations. Because such stimuli do not produce any ON–OFF pairings (Figure 1, middle), the model wrongly predicted no luster in these cases. For other stimulus conditions that produced ON–OFF pairings along with ON–ON and/or OFF–OFF pairings, the magnitude of the lustrous response was more or less underestimated by the model (an example stimulus is shown in Figure 1, right). As these cases indicate, a lustrous impression does not require the presence of ON–OFF pairings but can also be produced, or enhanced, by ON–ON and OFF–OFF pairings.

Georgeson, Wallis, Meese, and Baker (2016) addressed this issue in their study on binocular contrast discrimination. Their model for the detection of interocular contrast differences is based on two perceptual cues—namely, contrast and luster. The authors also assume that the luster cue is exclusively activated by stimuli that produce ON–OFF pairings. Regarding the lustrous appearances evoked by stimuli with equal contrast polarities, they found evidence that in these cases luster may result from noise at the level of the luminance signals. They used the same center–surround stimuli as in the study by Anstis (2000) and found that the luster judgments of Anstis' study could be predicted more accurately when noise was added to the luminances of the center patches and the surround element. Georgeson et al. (2016, p. 113) therefore assumed that "lustre is mainly induced by opposite signs of contrast, and that luminance noise can explain why lustre diffuses into the same-polarity quadrants."

In the present study, however, we propose an alternative approach. We expanded our original model (Wendt & Faul, 2020), which was exclusively based

on ON–OFF pairings, such that in the modified version also ON–ON and OFF–OFF pairings are included—albeit with considerably lower weights relative to the ON–OFF combinations. In order to test the predictive power of the model, we conducted a series of four experiments using different sets of center–ring–surround configurations as test stimuli. With the improved model, not only did we largely prevent the above-mentioned problems that occurred with the original version, but we also found a higher overall predictive power, with a proportion of explained variance of more than 80%.

Modified version of the interocular conflict model

Following the approach from Anstis (2000), binocular luster emerges from some kind of binocular conflict between monocular contrast signals that results from the activity of two different types of retinal ganglion cells. These ganglion cells have receptive fields with a circular symmetric center–surround structure (Schiller, 1992; Wienbar & Schwartz, 2018). One cell type, the ON-center cell, has an excitatory center area and an inhibitory surround and is therefore responsive to luminance patterns that stimulate the central area more strongly than the surround. The other cell type, the OFF-center cell, has an inhibitory center area and an excitatory surround and strongly responds to luminance patterns that stimulate the center area less than the surround. In our computational model, these monocular contrast detector systems are represented by an LoG filter kernel (Marr, 1982; Marr & Hildreth, 1980) with a certain radius r where the width of the Gaussian equals $r/2$ (note that the optimal filter radius will be determined empirically in our experiments). The circular center area of this LoG filter kernel is characterized by negative weights and the surround area by positive weights; that is, an ON signal resulting from an incremental stimulus pattern would be represented by a negative filter output and an OFF signal that results from a decremental luminance pattern by a positive filter output.

As a first step in the sequence of processes assumed in our model, a convolution with the LoG filter kernel is applied to both half-images of the dichoptic stimulus, I_l and I_r, using the LaplacianGaussianFilter function of Mathematica (see the second row in Figure 1). Note that, in the figure, the pixels in the two original images I_l and I_r represent luminances, not RGB values, and, in order to keep the pixel values within the interval of [0, 1], the original luminance values of the stimuli were divided by 100. Contrary to our former model, in which the entire stimulus was taken into account (Wendt & Faul, 2020), only those pixels of the resulting filtered images I'_l and I'_r are considered that belong to the target area, which is defined here as those pixels at which the two original half-images I_l and I_r differ in luminance, because only these parts of the stimulus will give rise to a lustrous sensation. In most of our experimental conditions, the target area coincides with the center patch area of the stimuli. However, in stimuli with a blurred ring as they were used in Experiment 3, also parts of the ring elements can have an interocular luminance difference and will therefore be included in the calculation.

In the next step, corresponding pixels in the two filtered images are analyzed with respect to the sign of their filter outputs. In our former model, only those corresponding pixels were taken into account that showed opposite signs and therefore represented contrast signals from an ON-center cell in one eye and an OFF-center cell in the other eye ($Sign[I'_l(x,y)] \neq Sign[I'_r(x,y)]$; see the yellow parts in the bottom row of Figure 1). However, as described in the Introduction, we modified our model such that corresponding pixels with equal signs are also considered to account for the fact that binocular luster can also be elicited by dichoptic center–surround stimuli with consistent contrast polarities—that is, by inc–inc or dec–dec stimuli that will always trigger contrast detector cells of the same type at corresponding retinal locations (Figure 1, middle). However, these consistent contrast polarity pairings will have different weights than the ON–OFF pairings, and, as we will see below, these weights will also differ between ON–ON and OFF–OFF pairings.

In the following step, the two monocular contrast values are combined into a measure of binocular conflict for each corresponding pixel pair within the target area. In order to account for rounding errors, we only consider filter value pairs where both (unsigned) monocular values exceed a fixed threshold of 10^{-6}. Because the mechanism underlying binocular luster responds to interocular contrast differences, a procedure based on binocular differencing (Henriksen & Read, 2016; Kingdom, 2012) is used: $w*Abs[I'_l(x,y) - I'_r(x,y)]$, where w represents the weight by which the full-wave rectified difference of the filter values is multiplied. This weight depends on the specific type of interocular contrast pairing (i.e., ON–OFF, ON–ON, or OFF–OFF). Note that, in the case of ON–OFF signal combinations, this procedure is equivalent to the summation of the absolute filter values. To determine the weights w for the three different types of contrast pairings, a grid search was performed on the total set of our empirical data (Appendix A). Because only relative weights are of interest, the weight for the ON–OFF signals was fixed to $w = 1$. We found that the best model predictions in terms of R^2 occurred when, for the ON–ON signals, a weight between 0.04 and 0.12 was used and, for the OFF–OFF signals, a weight was used

that was about half as high as the ON–ON weight. This difference in the weights for the two same-sign pairings is also in good agreement with the finding by Wendt and Faul (2019) that the strength of the lustrous response is weaker in dec–dec stimulus conditions (that exclusively produce OFF–OFF signals) compared with inc–inc stimuli (that exclusively trigger ON–ON mechanisms). For reasons that are explained in Appendix A, we decided to use the fixed set of weights ($w_{ON–OFF} = 1$, $w_{ON–ON} = 0.05$, and $w_{OFF–OFF} = 0.025$) for the three different contrast sign pairings.

The total sum of all binocular conflict values is used as a measure for the amount of interocular conflict caused by the entire dichoptic stimulus. Note that, in our former model, we generated three different conflict measures, all of which strongly correlated with the empirical luster data (Wendt & Faul, 2020). These were (1) the total number of pixel pairs in the stimulus representing ON–OFF pairings, (2) the total sum of binocular conflict signals produced by these ON–OFF pairings, and (3) the average amount of binocular conflict signals—that is, the total sum divided by the number of ON–OFF pixel pairs. In the modified version of our model, however, only the total amount of binocular conflict is used as output. This is due to the fact that the number of pixel pairs that are taken into account in the present model are almost constant across different stimulus conditions when same-sign contrast signals (ON–ON and OFF–OFF) are also included (Figure 1). In our former model, we only considered the yellow areas in the figure, whereas the modified version additionally includes the green and red areas, which together generally are the same size among the different stimulus conditions.

Another relevant aspect of our model, which will be examined in detail in the experimental sections, is the quantitative relationship between the perceived luster strength and the amount of interocular conflict. Some results of our former study imply a nonlinear relationship between these two measures which suggests a nonlinear conflict response function (see also the General Discussion).

Methods

Subjects

Five subjects participated in the experiments (three females and two males); one of them was an author of the study (GW). Their ages ranged from 22 to 49 years with a median of 22 years. All had normal or corrected-to-normal visual acuity. Prior to the experiments, the subjects gave their consent after being informed about their rights, the purpose and procedure of the study, and potential risks, following the basic tenets of the Declaration of Helsinki.

Apparatus

For the display of the stimuli, we used a 24-inch monitor with a resolution of 1920 × 1200 pixels (EIZO CG243W; Eizo, Inc., Ishikawa, Japan), which was calibrated according to a standard method (Brainard, 1989) using a JETI specbos 1211 spectroradiometer (JETI Technische Instrumente GmbH, Jena, Germany). The monocular half-images of the dichoptic stimuli were presented side-by-side on the screen and fused using a mirror stereoscope (ScreenScope; Stereo Aids, Albany, Western Australia, Australia) that was mounted on the monitor. During the experiment, the dichoptic test stimulus was always displayed together with either an anchor stimulus or a matching stimulus (see below). The vertical center-to-center distance between the test and comparison stimulus was 13.7 degrees of visual angle (dva). The viewing distance was 50 cm.

Test stimuli

In the present study, four experiments were conducted in which the dichoptic test stimuli were all based on square-shaped, center–ring–surround configurations. The side length of the center patches was 2 dva. As in our previous study (Wendt & Faul, 2020), we used square-shaped instead of circular stimulus elements in order to avoid the local occurrence of step-like pixel patterns or anti-aliasing effects. The surround element of the test stimulus filled the entire (upper or lower) half of the monitor, whereas in the other half either an anchor stimulus or a matching stimulus of the same type was presented. To facilitate fusion of the monocular half-images we used 1-pixel-thick right angles at each corner of the monocular center–ring elements with a side length of 4.7 dva and a distance of 8.6 dva between each corner point and the center of the stimuli. The color of these fusion locks was either black or white, depending on the surround luminance of the current stimulus pair.

In three of the four experiments, we investigated how the strength of the lustrous impression depends on certain spatial features of the ring element—namely, the width, the amount of dashing, and the blurriness of the ring (Figure 2). Luster judgments were obtained under two fixed levels of interocular luminance difference between corresponding center patches. In one level, this luminance difference was 20 cd/m^2 (one center patch set to 15 cd/m^2, the other to 35 cd/m^2); in the other level, the luminance difference was 30 cd/m^2 (with center patch luminances set to 10 cd/m^2 and 40 cd/m^2). The

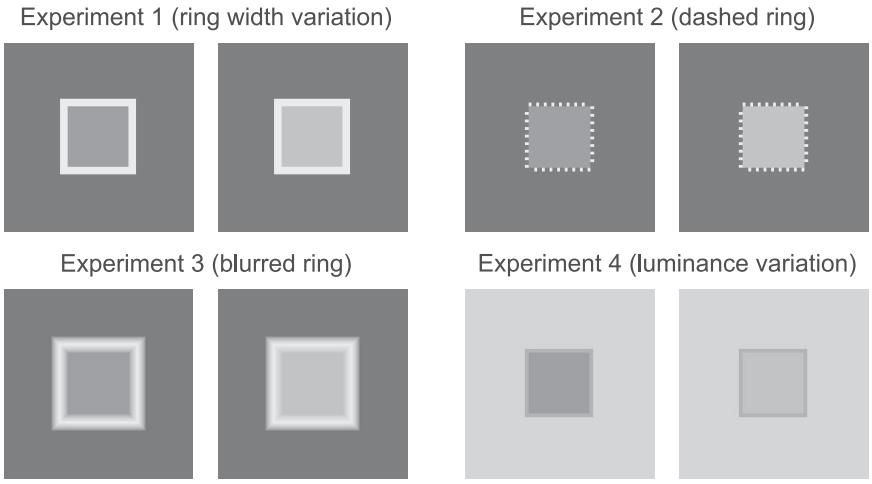

Figure 2. In the present study, we examined four different types of center–ring–surround stimuli separately in four different experiments. In Experiments 1 to 3, four different ring–surround luminance combinations were tested (see Figure 3; for the example stimuli, the condition *Consistent A* was used). Additionally, different spatial properties of the ring element were varied. In Experiment 1, the width of the ring was varied in 15 steps between 0 and 1.06 dva (the example shows a ring width of 0.25 dva). In Experiment 2, the ring width was kept constant with 0.125 dva while the dashing of the ring was varied. To this end, each side of the square shaped ring was subdivided in eight segments whose length was varied in eight steps between 1 and 8 pixels (the example shows ring segments with a length of 3 pixels). In Experiment 3, the ring was presented with a blurring that was controlled by varying the width of the ring (see Experiment 3 for details; in the present example, the ring is shown with a width of 17 pixels). In Experiment 4, the width of the ring element was kept fixed with 0.125 dva while the luminances of the ring and the surround element were systematically varied. In total, 256 ring–surround luminance combinations were tested in Experiment 4.

two different luminances of the center patches were swapped between eyes such that in half of the trials the left eye saw the brighter and the right eye saw the darker center patch, and vice versa. In this way, we aimed to compensate for possible effects of ocular dominance. The same procedure was applied to the matching stimulus (see below).

In addition, four different combinations of ring and surround luminance were tested in these experiments. In two of these conditions, the luminances of ring and surround were chosen such that both center patches always had consistent luminance contrast polarities with respect to the ring and surround elements (Figure 3). In the *Consistent A* condition, the center patches were always decremental to the ring luminance (which was set to 75 cd/m^2) and incremental to the surround luminance (5 cd/m^2), whereas in the *Consistent B* condition, the patch luminances were always incremental to the ring luminance (0 cd/m^2) and decremental to the surround luminance (50 cd/m^2). In the remaining two conditions, the patches showed reversed luminance contrast polarities with respect to either the ring or the surround element. In the *Reversed A* condition, the luminances of the center patches straddled the luminance of the surround (25 cd/m^2) but were both decremental to the ring luminance (75 cd/m^2), whereas in the *Reversed B* condition, the center patch luminances straddled the ring luminance (25 cd/m^2) but were incremental to the luminance of the surround (5 cd/m^2).

In the remaining experiment, we used a fixed spatial layout for the test stimuli where the ring had a width of 0.125 dva and systematically varied the luminance of both the ring and the surround element (see Experiment 4).

Matching stimulus

As matching stimulus, we used the same center–ring–surround configuration in all four experiments, with the luminance of the surround set to 10 cd/m^2. The ring was presented with a fixed width of 0.125 dva (= 4 pixels) and a luminance of 25 cd/m^2. The ring luminance was chosen such that the two center patches, whose mean luminance was kept fixed at 25 cd/m^2 and whose interocular luminance difference could be interactively manipulated by the subjects, always had reversed contrast polarities with respect to the ring. Under these luminance conditions, a large range of luster strengths could be produced. Furthermore, by the use of a ring configuration instead of a simple center–surround stimulus with reversed contrast polarities, the occurrence of rivalry effects could be significantly reduced (Wendt & Faul, 2020). The luminances

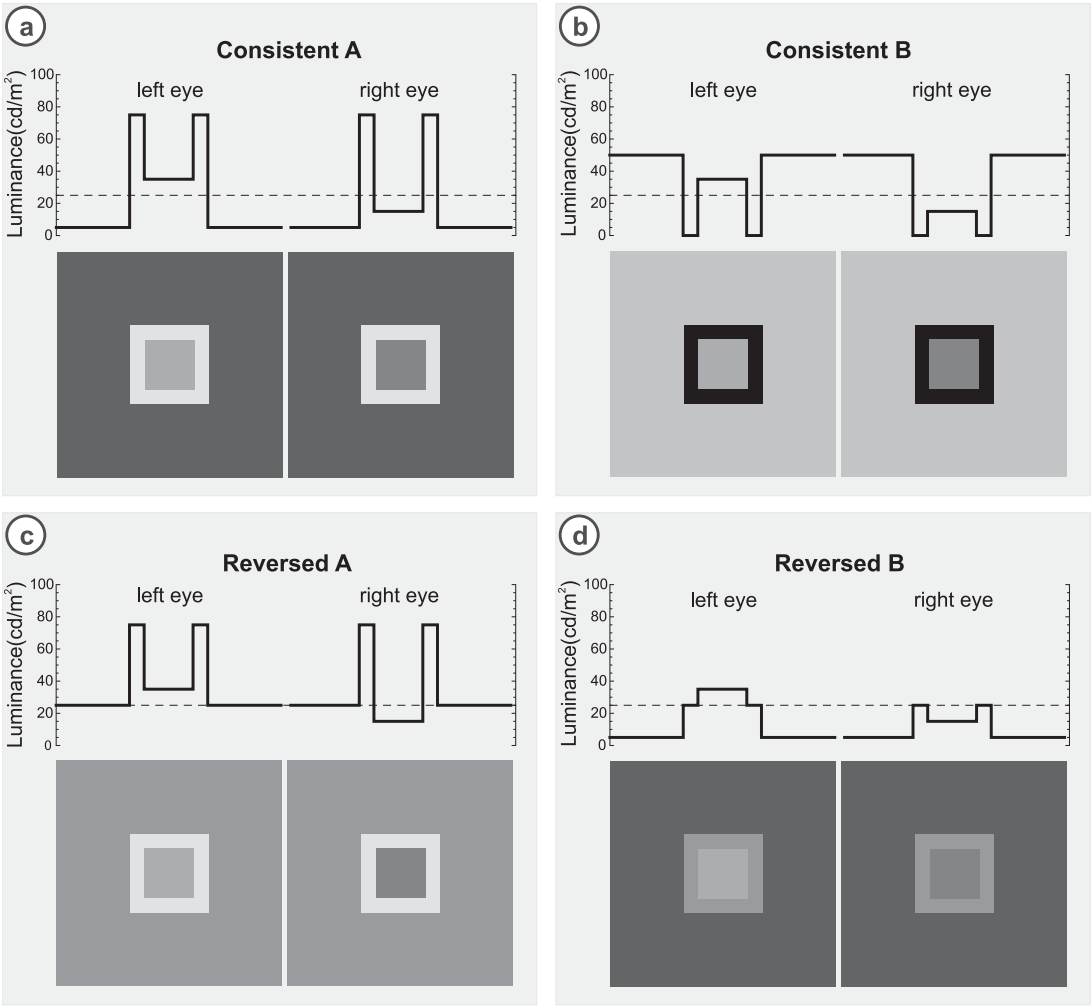

Figure 3. Schematic representation of the four different ring–surround luminance combinations as they were used in Experiments 1 to 3 (see also Figure 2 in Wendt & Faul, 2020). In the two conditions *Consistent A* and *Consistent B*, the contrast polarities of the center patches always had the same sign with respect to the ring and the surround luminances (a and b). In the two remaining conditions, the center patches had reversed contrast polarities with respect to either the surround (*Reversed A*) or the ring (*Reversed B*) but consistent polarities regarding the other stimulus element (c and d, respectively).

of the two monocular center patches ($C_{l,r}$) were calculated with the following equation:

$$C_{l,r} = 25\ \text{cd/m}^2 + 0.5 * (\alpha^k * 50\ \text{cd/m}^2),$$
$$\text{for } 0 \leq \alpha \leq 1 \text{ and } k = 1/0.77$$

where α represents the value set by the subjects using the left and right arrow keys of the keyboard. Exponent k was used to account for the finding that the strength of the lustrous impression is a decelerating function of the interocular luminance difference (Wendt & Faul, 2019), so that, with the given exponent k, constant changes in α would approximately produce perceptually equidistant changes in perceived luster strength.

Procedure

In each trial, the strength of the lustrous impression was measured using two different procedures. In the first part, a rating task was employed in which the subject had to rate the perceived strength of the luster on a scale between 0 (“matte”) and 5 (“maximally lustrous”). In order to facilitate the detection of rather subtle degrees of luster, the test stimulus was presented together with a matte anchor stimulus (e.g., Paillé, Monot, Dumont-Bècle, & Kemeny, 2001; Wendt & Faul, 2020). The ring and surround element of the anchor stimulus were always identical to those of the test stimulus; however, in order to produce a matte appearance in the target area of the anchor, the two monocular center patches were set to the same luminance of 22.5 cd/m^2.

In all cases where the test stimulus was rated with a value greater than zero, the rating task, after a 1-second lasting interval of dark adaptation, was followed by a matching task in which the same test stimulus was presented together with the matching stimulus (see the previous section). The task of the subjects was to adjust the interocular luminance difference between the two center patches of the matching stimulus such that the resulting lustrous appearance in the fused percept was indistinguishable from that induced by the test stimulus. In cases where the test stimulus was perceived as non-lustrous in the rating task, the matching task was skipped, and the setting was stored with an interocular luminance difference of 0 cd/m^2. The next trial started after a dark adaptation interval of 3 seconds.

No time constraints were imposed on the subjects for the two tasks. They were also asked to move their eyes over the stimuli during presentation, as it was found that fixation of the target area can lead to a weakening of the lustrous effect due to local adaptation (Kiesow, 1920; Sachsenweger, 1960; Wendt & Faul, 2022). For the entire experimental series, the subjects needed between 20 and 25 hours, divided into 10 to 12 blocks of about 2 hours each.

Experiment 1: Varying the width of the ring element

In our previous study (Wendt & Faul, 2020), we varied the ring width in our stimuli between 0 and 3.5 dva in steps of 0.5 dva; however, for ring widths larger than 1 dva, the corresponding luster judgments remained almost constant. In the present experiment, we therefore examined the interval of ring widths in more detail where stronger variations of the lustrous responses were to be expected (i.e., 0–1.06 dva). In total, 15 different ring widths in varying resolutions were examined. For the interval between 0 and 0.125 dva, the ring width was varied in steps of 1 pixel (= 0.03125 dva), for the interval between 0.125 and 0.4 dva in steps of 2 pixels and for the remaining ring widths in steps of 4 pixels. In addition, as stated in the Method section, luster judgments were obtained for four different ring–surround luminance combinations of the test stimuli (Figure 3) as well as for two different levels of the interocular luminance difference between corresponding center patches. Each of the 120 condition combinations (15 ring widths × 4 ring–surround luminance combinations × 2 levels of interocular luminance differences) was tested eight times, and the resulting 960 trials were presented in random order.

Results

Figure 4 shows the results of the ring width experiment for all five subjects (which are indicated by different colors). In the diagrams, the luster ratings and settings are shown as a function of the ring width, separately for the four different ring–surround luminance conditions (columns) and the two interocular luminance difference levels (labeled as "low" or "high"; see the upper and lower half of the panel).

All subjects generally show the same trends that were already evident in the previous study, albeit partially to a different extent. These trends are also comparable between the data obtained with the rating and the matching task (with correlation coefficients between $r = 0.911$ and $r = 0.971$ individually and $r = 0.976$ for the data averaged across all subjects). Due to the finer resolution of the ring widths in the present experiment, the range in which the variation takes place is more clearly visible. The changes seem to occur in a small range between about 0 and 0.4 dva, before the data curves flattens considerably.

In the two *Consistent* conditions, the addition of a ring of only 1 pixel width causes a dramatic increase in the strength of perceived luster in comparison with the no-ring condition. This luster level is largely maintained for subsequent ring widths up to about 0.2 dva before, for some subjects, the luster level then continuously flattens with increasing ring width. For other subjects, however, this level seems to be largely maintained over the remaining distance.

As expected, the two *Reversed* conditions show opposite trends. In the *Reversed A* condition, where the two center patches had reversed contrast polarities to the surround luminance but were both decremental to the ring luminance, the perceived glossiness first decreases dramatically with increasing ring width until the curves change more slowly from a ring width of about 0.2 dva before becoming nearly constant for ring widths larger than 0.4 dva. In the *Reversed B* condition, where the center patches have reversed contrast polarities only with respect to the ring luminance, the glossiness shows a rapid increase, which slowly levels off from about 0.2 dva and then also becomes a constant function from about 0.4 dva.

Model fit and discussion

The finding that the luster curves flatten above a certain ring width could already provide an indication of the size of the receptive fields of the underlying ON or OFF mechanisms as postulated by the model. Ring widths greater than about 0.4 dva do not seem to have any further effect on perceived luster, suggesting that larger ring widths exceed the range of the respective receptive fields.

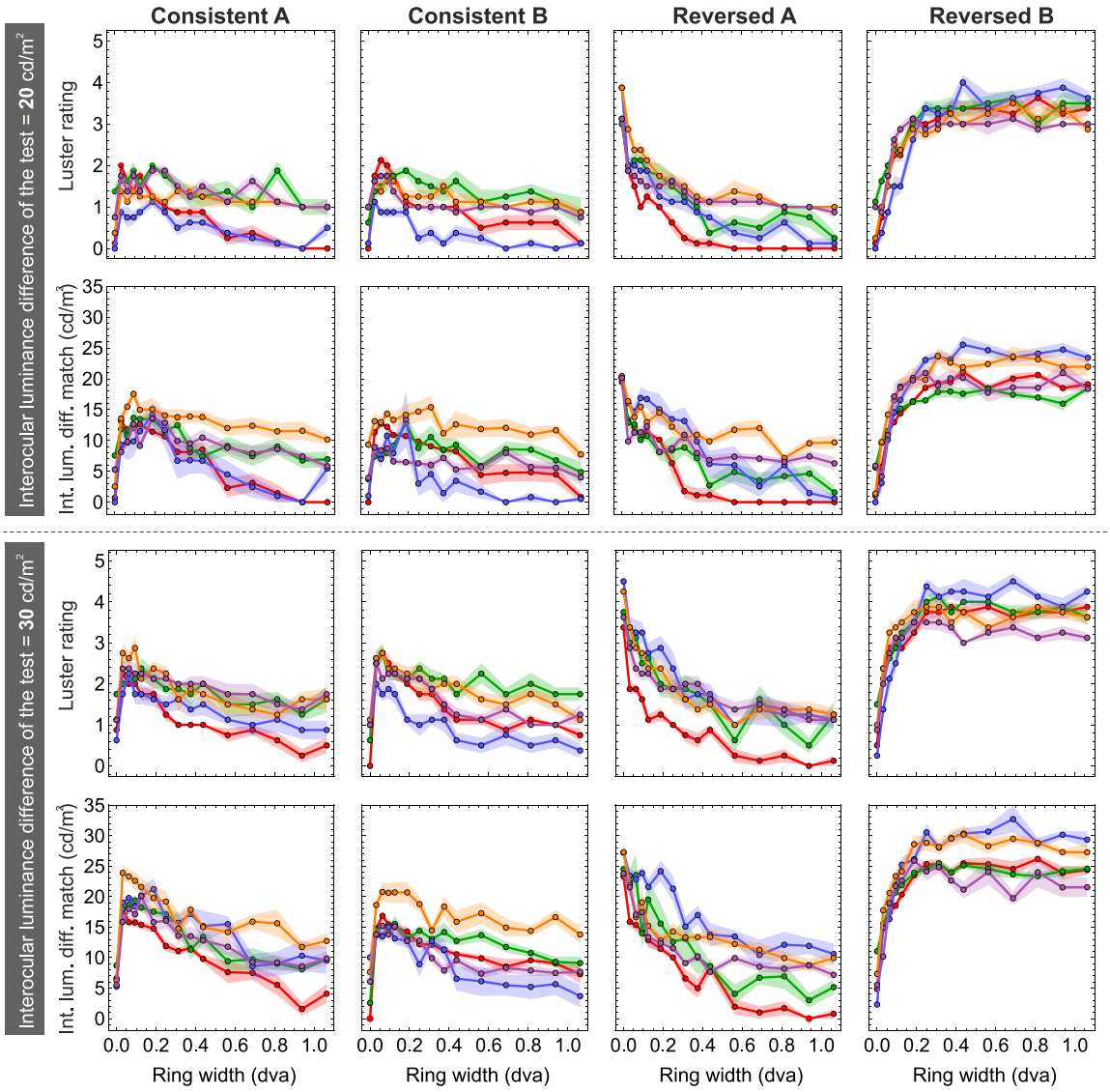

Figure 4. Results of Experiment 1 (ring width variation). The upper half shows the results for the low interocular luminance difference condition, and the lower half shows the results for the high luminance difference condition. The upper row in each half shows the results of the rating task, the lower row those of the matching task. For each of the four different conditions for the ring and surround luminance (columns), the diagram shows the mean rating or matching results as a function of the ring width for all five subjects (colored curves). The transparent areas represent 1 *SEM* in both directions.

This is supported by the model predictions calculated for different sizes of the filter kernel. Generally, we found that the relationship between the empirical luster settings and the amount of interocular conflict can be well described by a nonlinear function of the form $y = ax^c$, where y represents the mean luster settings, x the conflict measure (see the model description), and a and c the two parameters of the fit function. Because the amount of conflict depends on the size of the filter kernel, we searched for the LoG filter radius that produces the best fit for this function. As a measure for the goodness of fit we used the coefficient of determination R^2.

The top row in Figure 5 shows, separately for all subjects (S1–S5) and the pooled data across all subjects (S0), the R^2 values as a function of the radius of the LoG filter kernel, which was examined in the range between 5 and 30 pixels with a step size of 1 pixel. All curves show an inverted U-shaped relationship between filter size and the coefficient of determination R^2. The peak R^2 values for the averaged data with $R^2 = 0.886$ occurred at a filter radius of 15 pixels, which, for our experimental viewing conditions, is equivalent to a radius of 0.469 dva. For the individual datasets (S1–S5) the peak R^2 values range between 0.772 and 0.871, corresponding to filter radii between 13 and 17 pixels.

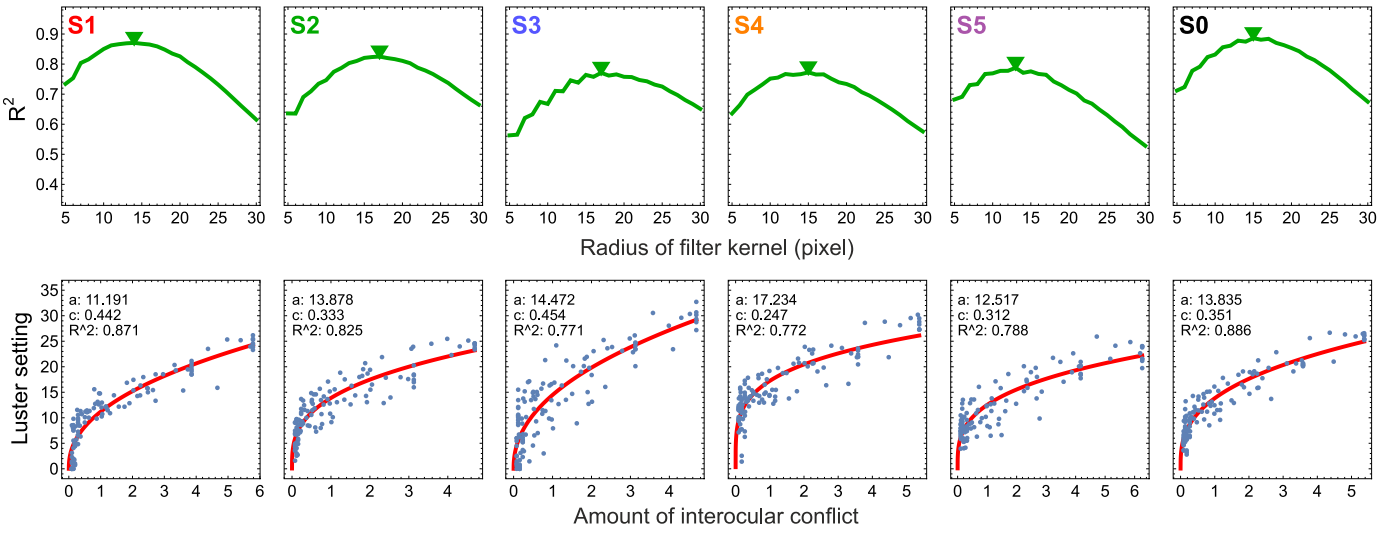

Figure 5. The amount of interocular conflict and luster judgments of Experiment 1 (ring width variation). In the upper row, the coefficients of determination (R^2) resulting from the fit of the empirical matching data with a nonlinear function of the conflict measures are shown in dependence on the radius of the LoG filter kernel, separately for the five subjects (S1–S5) and the averaged data (S0). The triangle in each diagram indicates the location of the peak R^2 value. In the bottom row, the mean luster settings are plotted against the corresponding conflict values as they were calculated using the filter size of the respective peak correlation (that is, at the triangle location of the upper diagram). The red curve shows the nonlinear fit function of the form $y = a*x^c$ (see the inlets in each diagram).

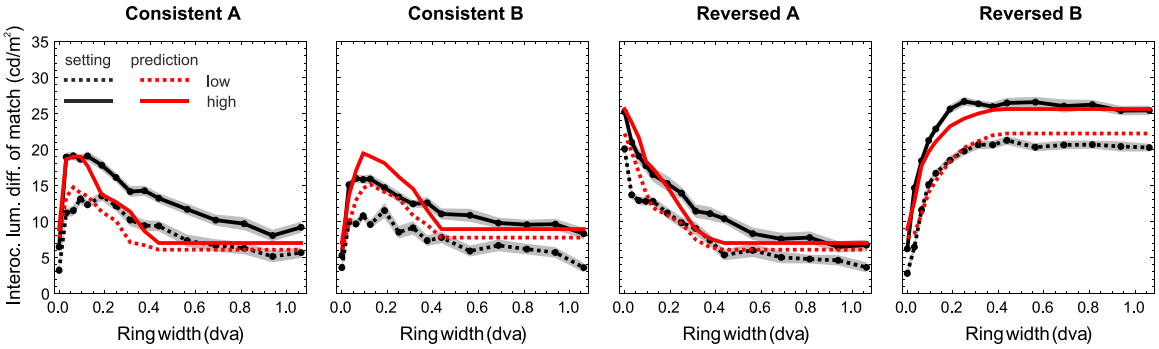

Figure 6. A direct comparison between luster settings and model predictions of Experiment 1 (ring width variation). The empirical luster settings averaged across all five subjects (black curves) are shown together with corresponding prediction lines (red curves) as they were calculated based on corresponding conflict values using a filter radius of 15 pixels and with the nonlinearity taken into account (see bottom right diagram of Figure 5). The columns correspond to the four different conditions for the ring and surround luminance. The low interocular luminance difference condition is indicated by dotted lines, and the high interocular difference condition is indicated by solid lines. The transparent areas around the black lines represent the *SEM* in both directions.

Note that the estimated size of the receptive fields of about 0.469 dva radius is somewhat smaller than the radius of about 0.625 dva found in our previous study (Wendt & Faul, 2020), a discrepancy that is likely due to the larger step sizes used in the former study with respect to both the ring widths tested and the radius of the filter kernel used in the modeling. In the bottom row of Figure 5, the empirical luster settings are shown in dependence on the corresponding conflict values, calculated with the respective filter size of the peak R^2 value, as indicated in the top row with a triangle. The red line shows the resulting fitting curve (with parameter values as displayed within the diagrams). As can be seen, there is a strong nonlinear relationship between luster settings and conflict values, which seems to be well fitted with a decelerating power function that is rather steep at lower model values and compressive at larger values (see also the General Discussion section).

Generally, the model describes the empirical luster judgments very accurately. This can also be seen in Figure 6 where the mean luster settings averaged

across all five subjects (black curves) are shown together with the corresponding prediction lines (red curves) as they were calculated by the model with the nonlinearity taken into account, with the parameter values given in the bottom right diagram of Figure 5—that is, *prediction value* = *a**(*model output*)^*c*. Note that, contrary to our initial, simpler model introduced in Wendt and Faul (2020), the present model also accounts for those weaker lustrous sensations that are produced by stimuli that do not lead to any ON–OFF pairings but exclusively to ON–ON or OFF–OFF pairings—cases for which our former model predicted no luster at all (Appendix A).

Experiment 2: Varying the dashing of the ring element

In this experiment, we used center–ring–surround stimuli in which the ring element was presented with different degrees of dashing. Each edge of the square-shaped ring was subdivided into eight segments, each a maximum length of eight pixels. The length of each segment was varied in eight steps from 1 to 8 pixels (0.03125–0.25 dva) (Figure 2). The width of the segments was kept constant at 4 pixels (0.125 dva). We used the same four luminance combinations for the

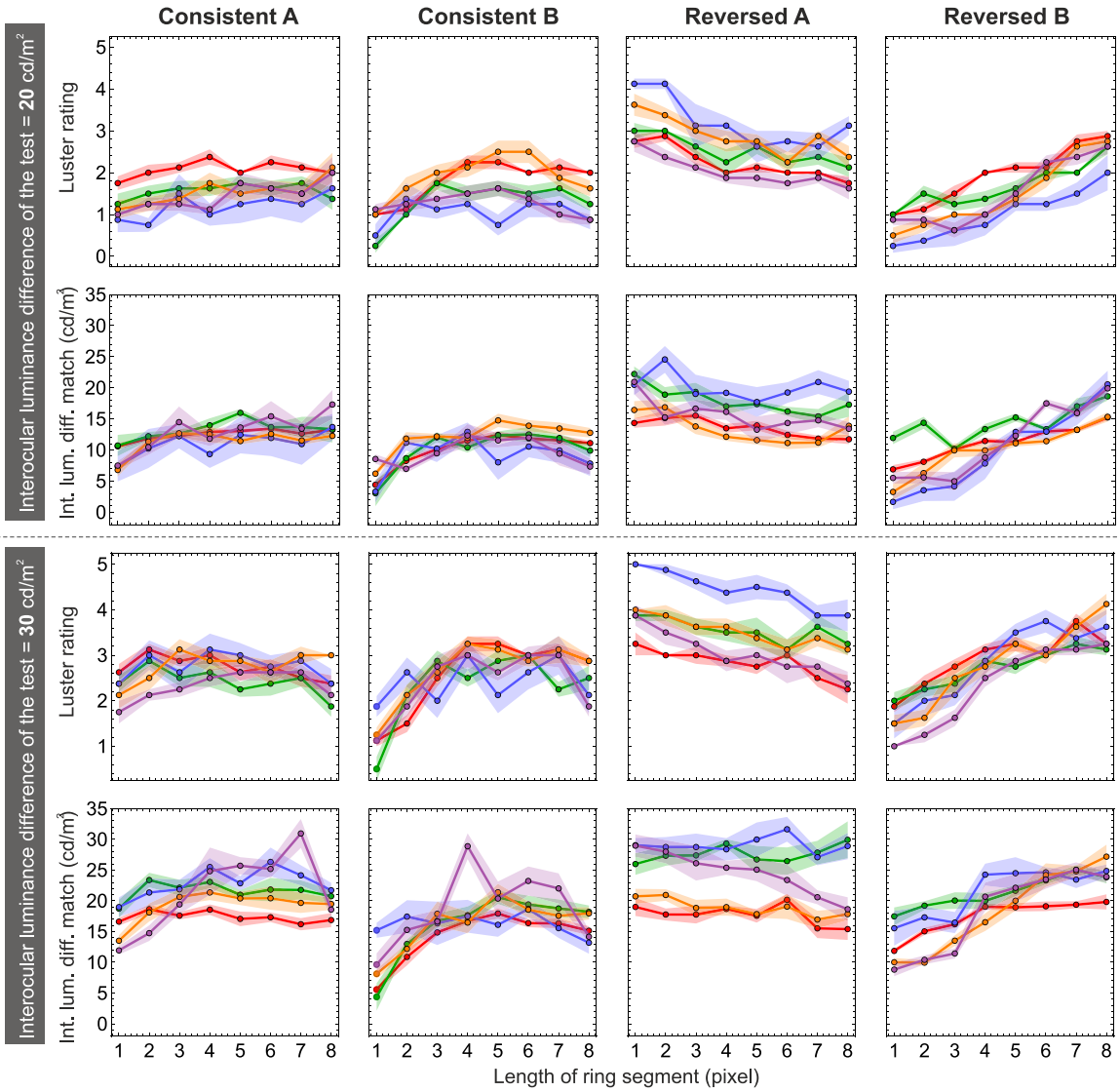

Figure 7. Results of Experiment 2 (dashed rings). The upper half shows the results for the low interocular luminance difference condition are shown, and the lower half shows the results for the high luminance difference condition. The upper row in each half shows the results of the rating experiment, the lower row those of the matching task. The columns give the results for one of the four conditions for the ring and surround luminance. Each diagram shows the mean rating or matching results as a function of the length of the ring segment for all five subjects (colored curves). The transparent areas represent the *SEM* in both directions.

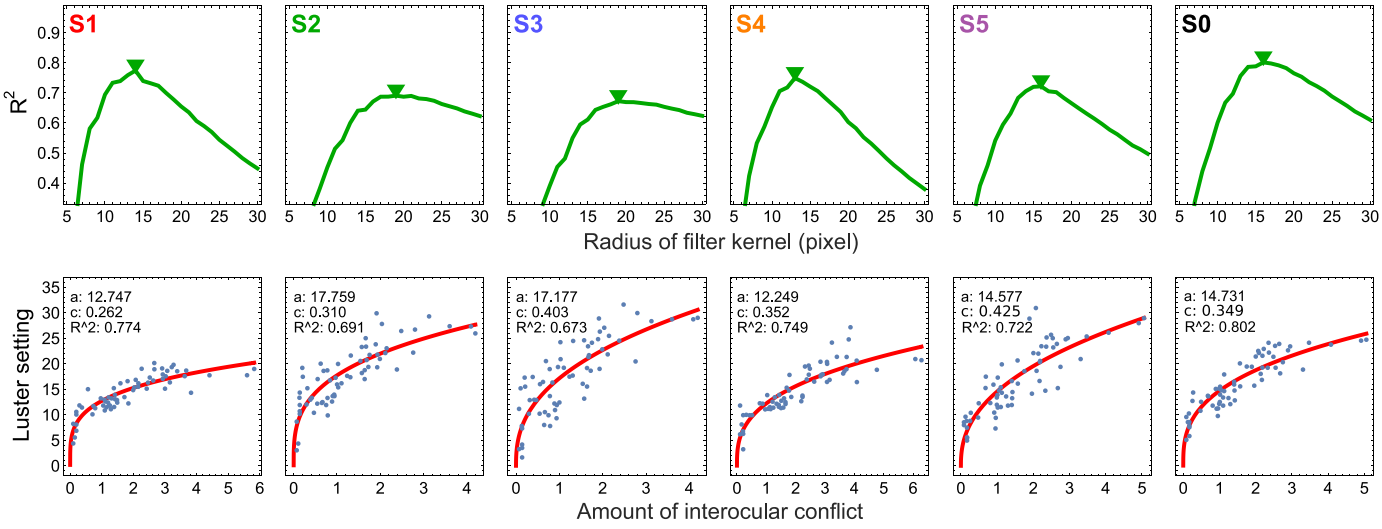

Figure 8. The amount of interocular conflict and luster judgments of Experiment 2 (dashed rings). In the upper row, the coefficients of determination (R^2) are shown in dependence on the radius of the LoG filter kernel, separately for the five subjects (S1–S5) and the averaged data (S0). The triangle in each diagram indicates the location of the peak R^2 value. In the bottom row, the mean luster settings are plotted against the corresponding conflict values as they were calculated using the filter size of the respective peak correlation (that is, at the triangle location of the upper diagram). The red curve shows the nonlinear fit function of the form $y = a*x^c$ (see the inlets in each diagram).

ring and surround element (Figure 2) and the same two interocular luminance difference conditions for the center patches of the test stimuli as in the previous experiment. All other aspects of the experimental setting were also identical. Each of the resulting 64 stimulus conditions (8 levels of ring dashing × 4 different ring–surround luminance combinations × 2 levels of interocular luminance difference between the center patches) was tested eight times in random order.

Results

In the same way as in Figure 4, Figure 7 shows the results of the dashed ring experiment. Each diagram is for one of the four different ring–surround luminance conditions (columns) combined with one of the two levels of interocular luminance differences (upper and lower half of the panels), and one of the two methods used to measure the strength of perceived luster. Each diagram shows the luster judgments as a function of the length of the ring segments, separately for all five subjects (colored curves).

Again, there is a strong correlation between the luster ratings and luster settings, with individual correlation coefficients ranging from $r = 0.859$ to $r = 0.938$ and a correlation $r = 0.96$ for the pooled data. Also the data trends between subjects are very similar: In the two *Consistent* conditions (leftmost columns in Figure 7), a monotonic increase of the luster judgments with increasing length of the ring segments up to a length of 4 pixels takes place, after which the luster curves seem to remain nearly constant. The initial increase of the perceived luster strength, however, is much more pronounced in the *Consistent B* than in the *Consistent A* condition. The two *Reversed* conditions again show opposite trends. Whereas in the *Reversed A* condition the luster curves show a slight decrease, the respective luster curves in the *Reversed B* condition strongly increase with increasing length of the ring segments.

Model fit and discussion

As in the previous experiment, we fitted the relationship between the empirical luster settings and conflict measures with a nonlinear function and determined the corresponding R^2 value. The fits were done for filter sizes between 5 and 30 pixels.

As can be seen in the top row of Figure 8, the peak R^2 values occurred at filter sizes between 14 and 19 pixels for the individual datasets (S1–S5) and at a filter size of 16 pixels for the averaged data (S0). The peak R^2 values varied between 0.673 and 0.774 for datasets S1 to S5, and a value of 0.802 was found for the averaged data (S0). Again, the bottom row shows the relationship between mean luster settings and the amount of interocular conflict for the filter size that led to the respective peak R^2 value. As in Experiment 1, it seems that the relationship between empirical luster

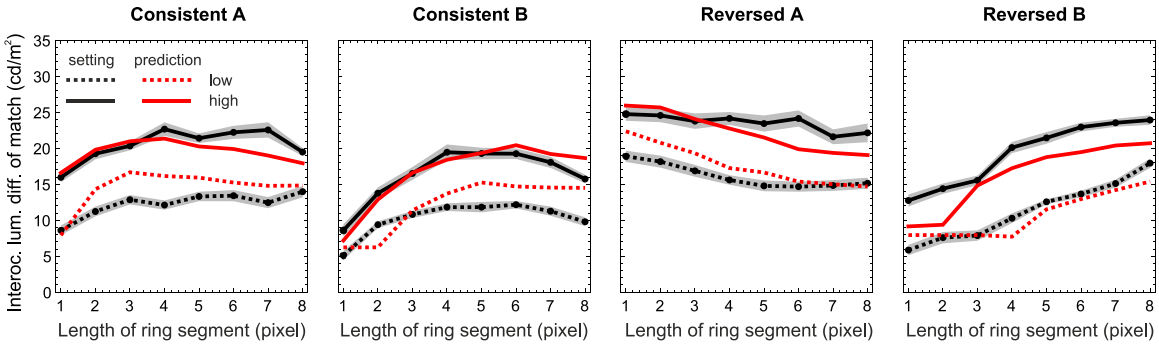

Figure 9. A direct comparison between luster settings and model predictions of Experiment 2 (dashed rings). Empirical luster settings averaged across all five subjects (black curves) are shown together with corresponding prediction lines (red curves) for the four different combinations of ring and surround luminance (columns). The low interocular luminance difference condition is indicated by dotted lines, and the high interocular difference condition is indicated by solid lines. The transparent areas around the black lines represent the *SEM* in both directions. Predictions were calculated from the corresponding conflict values using a filter radius of 16 pixels and taking the nonlinear response function into account (see bottom right diagram of Figure 8).

settings and conflict measures can be well fitted with a decelerating power function. Figure 9 compares the mean empirical luster settings (black curves) and corresponding predictions (red curves) for the different segment lengths realized in Experiment 2. Here, the same parameter values as in the bottom right diagram of Figure 8 are used.

Experiment 3: Varying the blurriness of the ring element

In Experiment 3, the blurriness of the ring element was varied. All other aspects were identical to those realized in the two previous experiments. In particular, we used the same four combinations for the luminances of the ring and the surround element (Figure 3) and the same two levels for the interocular luminance difference between the two center patches of the test stimuli. The blurring of the ring depended on the width w of the ring element which was varied in eight steps with two different resolutions: from $w = 1$ to $w = 7$ pixels in steps of 2 pixels and from $w = 11$ to $w = 23$ pixels in steps of 4 pixels. That is, each center patch was surrounded by w adjacent square-shaped rings of 1 pixel width, whose luminances were determined as follows: The middle ring element at $w_c = (w + 1)/2$ was set to the original ring luminance (L_R) of the respective condition shown in Figure 3. The luminance of ring k (L_k), with ring 1 adjacent to the surround, was set to $L_k = aL_X + (1 - a)L_R$, where $a = |k - w_c|/w_c$, $L_X = L_S$ for $k < w_c$, and $L_X = L_C$ for $k > w_c$. That is, the luminances of the ring elements between the middle ring element and the center patch were linearly interpolated between the luminance of the center patch (L_C) and the luminance of the ring (L_R). Similarly, the ring elements between the middle ring element and the surround were set to a luminance resulting from an interpolation between L_R and the surround luminance L_S (Figure 2). The 64 different condition combinations (8 levels of ring blurriness × 4 different ring–surround luminance combinations × 2 levels of interocular luminance difference between the center patches) were tested eight times each and presented in random order.

Results

The results of the blurred ring experiment are shown in Figure 10. In the diagrams, the luster judgments are shown as a function of ring width w (i.e., as a function of the degree of ring blurring). The correlation between the luster ratings and luster settings is on average slightly weaker than in the previous experiments, with individual correlation coefficients between $r = 0.686$ and $r = 0.875$, and $r = 0.931$ for the averaged data. In the two *Consistent* conditions, the luster judgments do not seem to vary much in dependence on the blurring. Generally, a small increase in perceived luster strength takes place within ring widths up to about 5 pixels, after which the curves remain largely constant. Opposite trends again occurred between the two *Reversed* conditions. In the *Reversed A* condition, luster judgments slightly decrease with increasing blurring up to a ring width of about 11 pixels and then flatten into a constant function. In the *Reversed B* condition, a comparatively steep increase in the luster strength takes place roughly within the same interval of ring widths which then also flattens for higher ring widths.

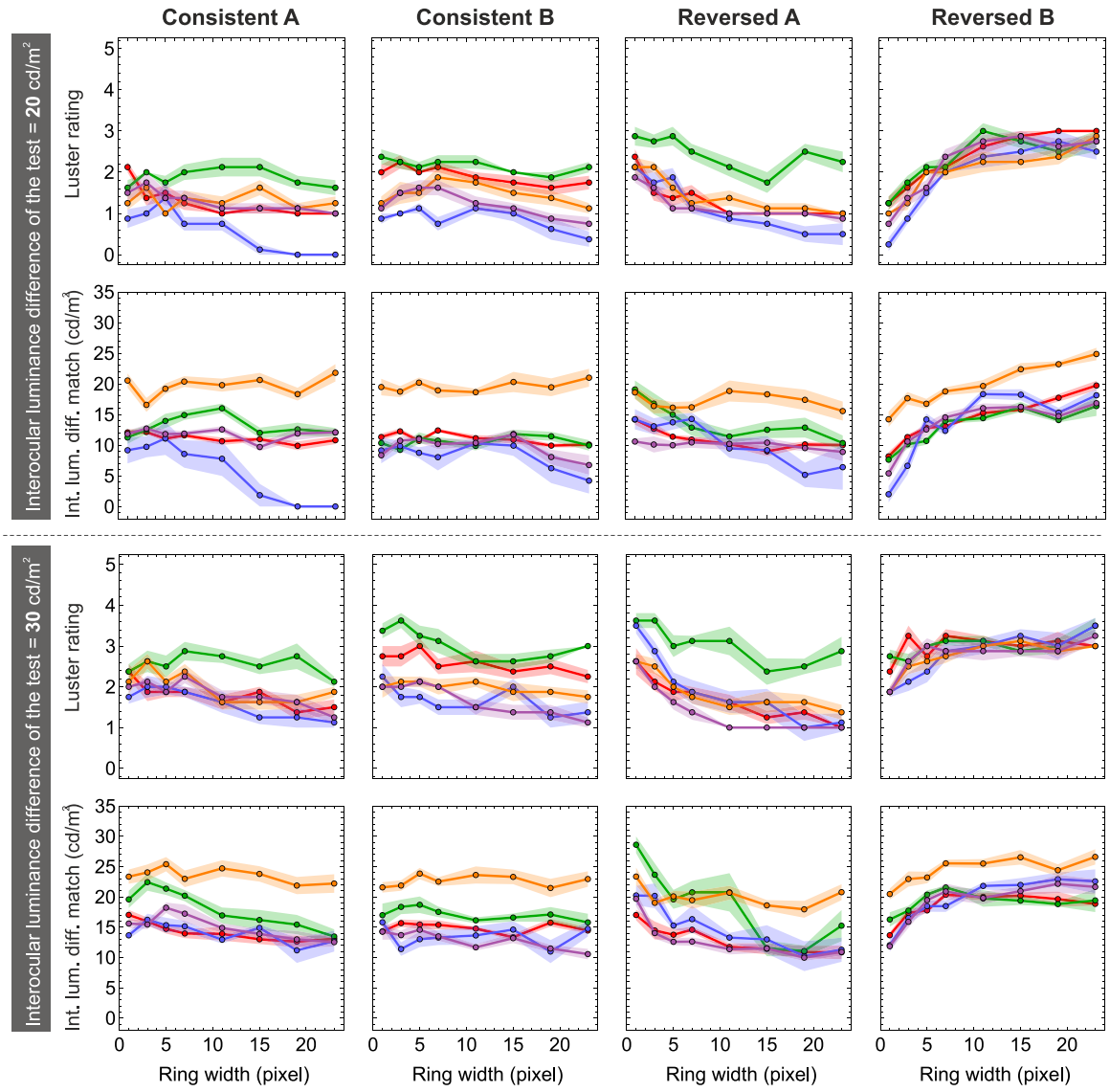

Figure 10. Results of Experiment 3 (blurred rings). The upper half shows the results for the low interocular luminance difference condition, and the lower half shows the results for the high luminance difference condition. The upper row in each half shows the results of the rating experiment, the lower row shows those of the matching task. Separated by the four different conditions for the ring and surround luminance (columns), each diagram shows the mean rating or matching results as a function of the width of the blurred ring element for all five subjects (colored curves). The transparent areas represent the *SEM* in both directions.

Model fit and discussion

As before, the top row in Figure 11 shows the coefficient of determination (R^2) in dependence on the LoG filter radius, separately for the five subjects (S1–S5), as well as for the averaged data (S0). Again, a high proportion of explained variance was found for the pooled data ($R^2 = 0.825$), but the R^2 values for the individual datasets were generally lower and varied considerably between $R^2 = 0.44$ (for subject S4) and $R^2 = 0.81$ (for subject S1). The filter sizes at which the peak R^2 values occurred were comparable to those found in the previous experiments. The filter size was 14 pixels for the pooled dataset (S0) and between 13 and 17 pixels for the individual datasets S1 to S5. Once more, it seems that the relationship between the empirical luster judgments and interocular conflict values can be well described by a decelerating power function. For the pooled data with optimal filter size, Figure 12 provides a direct comparison between the empirical luster settings (black curves) and corresponding predictions (red curves) depending on the amount of blurriness.

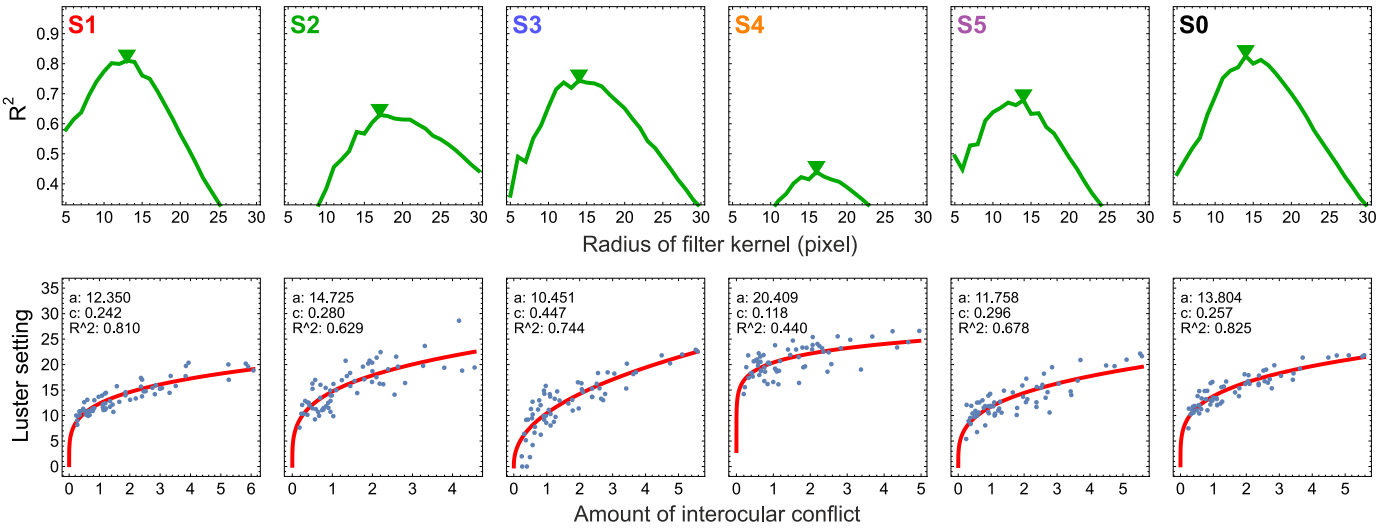

Figure 11. The amount of interocular conflict and luster judgments of Experiment 3 (blurred rings). In the upper row, the coefficients of determination (R^2) are shown in dependence on the radius of the LoG filter kernel, separately for the five subjects (S1–S5) and the averaged data (S0). The triangle in each diagram indicates the location of the peak R^2 value. In the bottom row, the mean luster settings are plotted against the corresponding conflict values as they were calculated using the filter size of the respective peak correlation (that is, at the triangle location of the upper diagram). The red curve shows the nonlinear fit function of the form $y = a*x^c$ (see the inlets in each diagram).

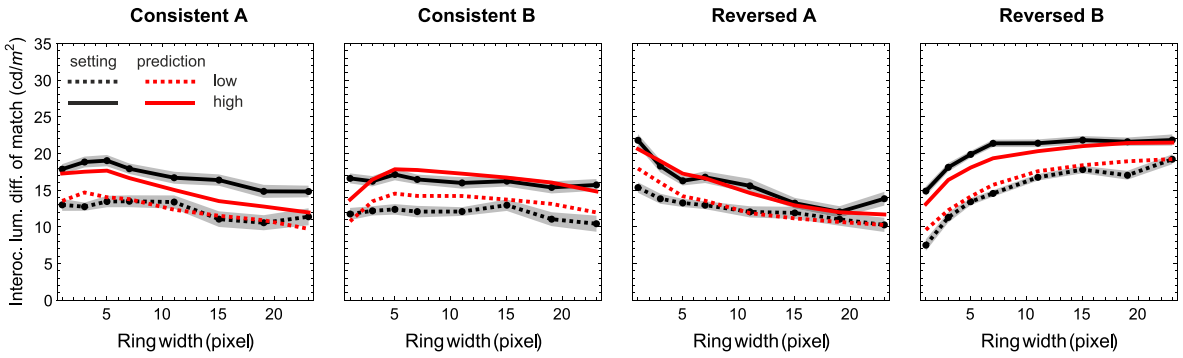

Figure 12. A direct comparison between luster settings and model predictions of Experiment 3 (blurred rings) for different ring widths that determine blurriness. The empirical luster settings averaged across all five subjects (black curves) are shown together with corresponding prediction lines (red curves) as they were calculated based on corresponding conflict values using an LoG filter radius of 14 pixels and with the nonlinearity taken into account (see bottom right diagram of Figure 11), separately for four different conditions for the ring and surround luminance (columns). The low interocular luminance difference condition is indicated by dotted lines, and the high interocular difference condition is indicated by solid lines. The transparent areas around the black lines represent the *SEM* in both directions.

Experiment 4: Varying the luminance of the ring and surround element

In the last experiment of the series, we examined test stimuli with a fixed spatial layout and systematically varied the luminance of the ring and the surround element independently from each other. The ring was presented with a width of 0.125 dva. The luminance of the ring and the surround were each varied in 16 steps between 0 and 100 cd/m^2. The interval between 0 and 50 cd/m^2 was varied in steps of 5 cd/m^2, and the interval between 50 and 100 cd/m^2 in steps of 10 cd/m^2. The interocular luminance difference between the center patches of the test stimulus was fixed at 20 cd/m^2, with the luminance of one patch set to 15 cd/m^2 and the other to 35 cd/m^2. Each of the 256 condition combinations (16 ring

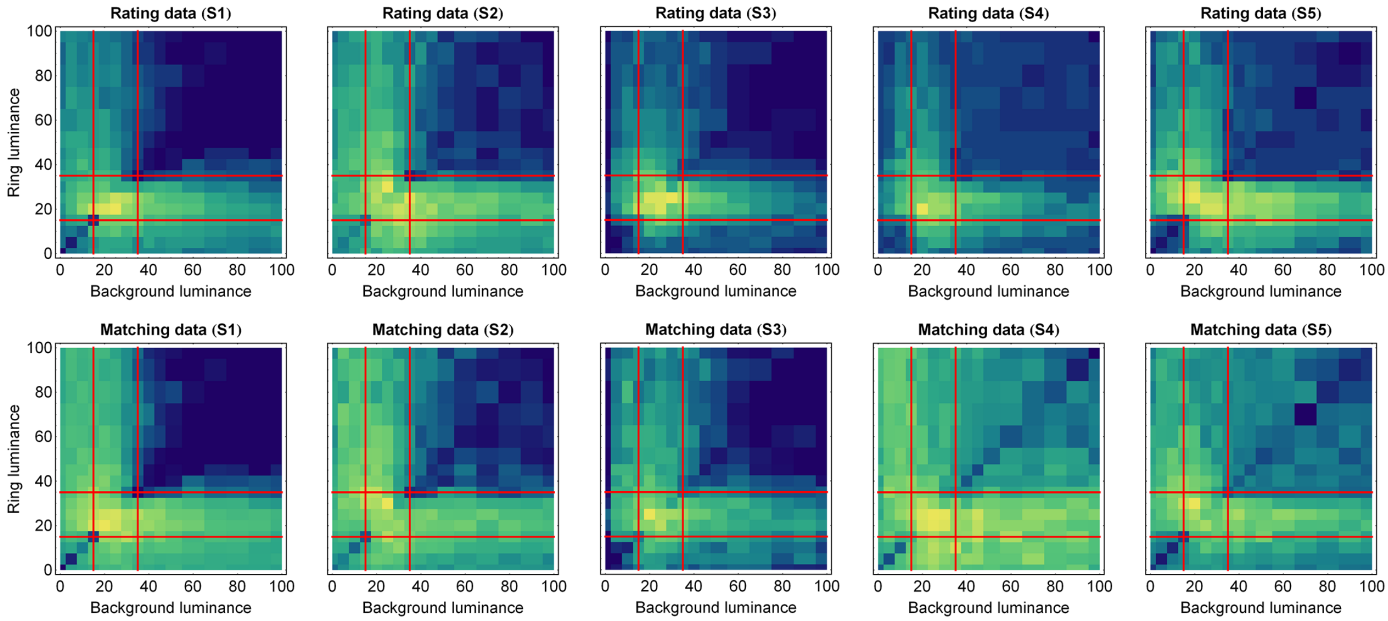

Figure 13. Results of Experiment 4 (luminance variation). For each of the five subjects (columns), density plots show the data of the rating (top row) and the matching task (bottom row). In each cell in the diagrams, the mean rating or matching value for the corresponding combination of background and ring luminance of the test stimulus is represented by a color on a continuous scale ranging from dark blue (no luster) to light yellow (strongest luster). The red lines indicate the fixed luminances of the two center patches.

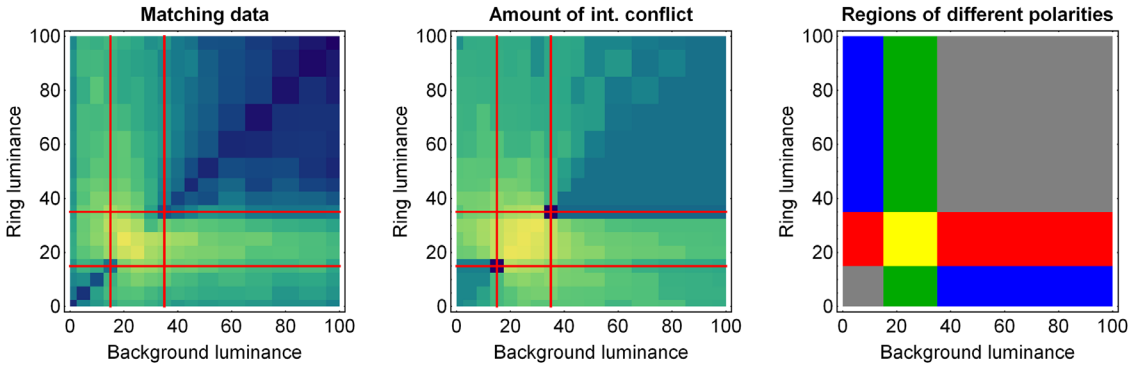

Figure 14. (Left) The density plot shows the matching results of Experiment 4 (luminance variation) averaged across all five subjects (cf. Figure 13). (Middle) Model predictions calculated based on corresponding conflict values at a filter radius of 11 pixels and taking the nonlinearity shown in the bottom right diagram in Figure 15 into account. (Right) Color scheme illustrating different types of interocular contrast polarities in regions of the plot (see text).

luminances × 16 surround luminances) was tested eight times.

Results

The density plots in Figure 13 show the results of the rating (top row) and the matching task (bottom row) separately for the five subjects (columns). The left side of Figure 14 shows the results for the matching task averaged across all subjects. In each diagram, the mean rating or setting in dependence on the ring and surround luminance of the test stimulus is represented by a color on a continuous color scale ranging from dark blue ("no luster") to light yellow ("maximal luster"). The red lines indicate the luminances of the two center patches of the test. Again, the rating and matching data were strongly correlated, with a correlation coefficient of $r = 0.96$ for the averaged data and individual coefficients between $r = 0.76$ (see the fourth column in Figure 13) and $r = 0.961$.

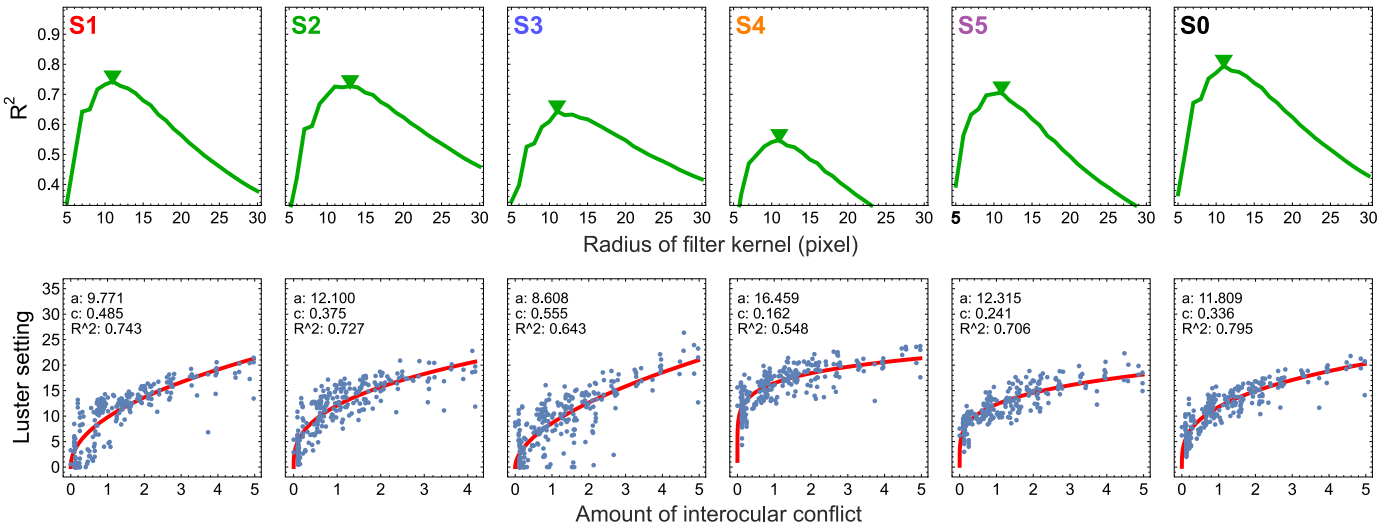

Figure 15. The amount of interocular conflict and luster judgments of Experiment 4 (luminance variation). In the upper row, R^2 values are shown in dependence on the radius of the LoG filter kernel separately for the five subjects (S1–S5) and the averaged data (S0). The triangle in each diagram indicates the location of the peak R^2 value. In the bottom row, the mean luster settings are plotted against the corresponding conflict values as they were calculated using the filter size of the respective peak correlation (that is, at the triangle location of the upper diagram). The red curve shows the nonlinear fit function of the form $y = a^*x^c$ (see the inlets in each diagram).

The red lines subdivide the diagrams into different sections that are associated with different dichoptic contrast polarities (Figure 14, right). The green and red sections represent stimulus conditions in which the two test patches have reversed contrast polarities with respect to either the background luminance (green) or the ring luminance (red). The yellow region in Figure 14 represents the conditions in which the center patches have reversed contrast polarities with respect to both the ring and the surround luminance. These three conditions were expected to produce comparatively strong sensations of luster, which is exactly what the empirical results show. The blue regions on the right side in Figure 14 represent stimulus conditions in which the center patch luminances are incremental to the ring and decremental to the surround, or vice versa. These conditions are comparable to the two *Consistent* conditions depicted in Figure 3. In the ring width experiment (see Experiment 1), comparatively strong lustrous sensations were measured for a ring width of 0.125 dva that was used in the present experiment (see the respective diagrams in Figure 4). Thus, this finding is also confirmed here. The two remaining regions (gray areas in Figure 14, right) represent conditions in which both center patches are either incremental or decremental to the luminance of both the ring and surround element. Here, the weakest lustrous sensations were to be expected which is generally confirmed by the empirical data, although there are comparatively strong interindividual differences in this respect.

Model fit and discussion

Again, the R^2 values are plotted as a function of the radius of the LoG filter kernel in the top row of Figure 15, separately for the individual datasets (S1–S5) and for the data averaged across all subjects (S0). The peak R^2 values were found at filter sizes that were slightly lower than those found in the other experiments. The filter size varied between 11 and 13 pixels for the individual sets and was 11 pixels for the pooled data (see the location of the green triangles in the top diagrams of Figure 15). As can be seen in the bottom row of Figure 15, the relationship between luster settings and conflict values (as calculated with the filter size at the respective peak R^2 values) again seems to be well described by a power function of the form $y = a^*x^c$. The goodness of fit in terms of R^2 ranged from 0.548 (see S4) to 0.743 for the individual datasets and was 0.795 for the averaged data. For direct comparison, the middle panel of Figure 14 shows the empirical luster settings and the model predictions in the same format. In this diagram, the predicted values were calculated using filter size 11 at the peak R^2 value of the pooled data, also taking into account the nonlinearity as shown in the bottom right diagram of Figure 15. Overall, the data profile predicted by the model is very similar to the one observed in the experiment. The largest deviations are found in the "same polarity" conditions, where the weak lustrous sensations observed in the experiment are somewhat overestimated by the model (cf. left and middle diagrams in Figure 14 in

the areas marked "gray" in the rightmost diagram in Figure 14).

General discussion

Putting it all together

In this study, we conducted four experiments to investigate how the strength of perceived luster depends on a variety of different dichoptic center–ring–surround stimulus conditions and how well the lustrous responses can be described by our model. Averaged across all subjects, we found a high predictive quality of our model, with proportions of explained variance in the four experiments ranging from 79.5% to 88.6%. This also holds when the data from all experiments are combined: Figure 16 shows the relationship between the empirical matching data and the corresponding conflict values for all 504 stimulus conditions tested in our experiments. In the top row of Figure 16, R^2 values resulting from fitting this relationship with a power function are plotted against the LoG filter radius. The peak R^2 values varied between 0.555 and 0.777 for the individual datasets (S1–S5) and was 0.818 for the averaged data (S0). They occurred at filter radii between 11 and 15 pixels for the individual datasets and at a size of 14 pixels radius for the pooled data. In the diagrams in the bottom row of Figure 16, the luster settings are plotted against the degree of interocular conflict. The nonlinear fitting curves (red) were again calculated using the filter sizes corresponding to the individual peak R^2 positions. Overall, the trends in the combined data are in good agreement with those found in the single experiments.

Differences to our original model

In our original model (Wendt & Faul, 2020), we considered only interocular contrast signals of opposite sign (that is, ON–OFF signal pairings; see the yellow parts in Figure 1). Stronger lustrous sensations could be well predicted by this model. Dichoptic stimuli with consistent between-eye polarities—stimuli that will always trigger the same type of contrast detector cell at corresponding retinal locations (see, for example, Figure 1, middle)—were ignored. Correspondingly, our former model (wrongly) predicted "no luster" in these cases (Appendix A), although such stimuli also produce weak lustrous responses. In our improved model, this problem could be solved by also including ON–ON and OFF–OFF signal pairings, albeit with a much lower weight than the ON–OFF pairings. The occurrence of ON–OFF pairings therefore still provides the main basis for stronger lustrous sensations (Anstis, 2000; Georgeson et al., 2016; Wendt & Faul, 2019).

An interesting question is how these weightings specifically affect the luster predictions obtained with simple center–surround stimuli. The 256 stimulus conditions used in Experiment 4 included some cases where the surround and the ring had the same luminance (see the diagonal in the diagrams shown

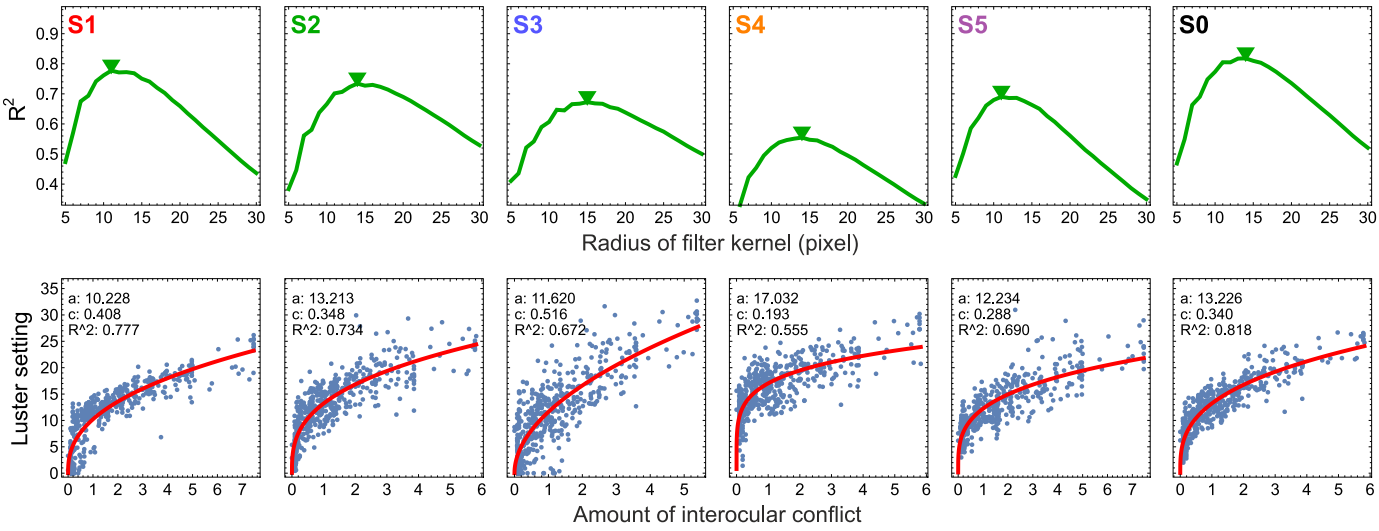

Figure 16. The amount of interocular conflict and luster judgments of all four experiments combined. In the upper row, R^2 values are shown in dependence on the radius of the LoG filter kernel separately for the five subjects (S1–S5) and the averaged data (S0). The triangle in each diagram indicates the location of the peak R^2 value. In the bottom row, the mean luster settings are plotted against the corresponding conflict values. The conflict values were calculated using the filter size of the respective peak correlation (that is, at the triangle location of the upper diagram). The red curve shows the nonlinear fit function of the form $y = a*x^c$ (see the inlets in each diagram).

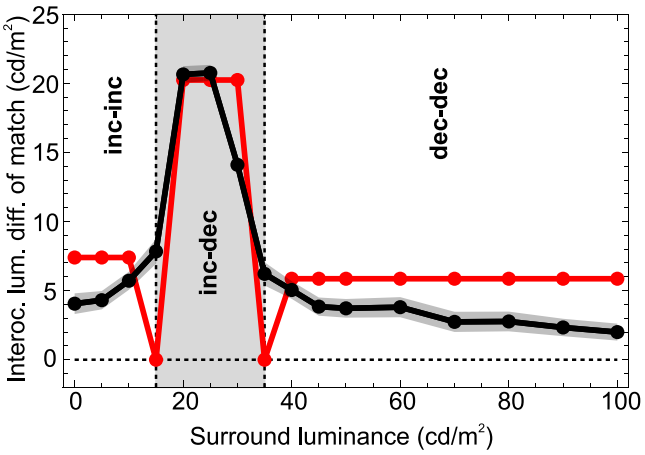

Figure 17. The mean luster settings (black curve; transparent area represents the *SEM*) and corresponding model predictions (red curve) as a function of the surround luminance for the subset of simple center–surround configurations of the stimuli used in Experiment 4 (see the diagonal in the diagrams of Figures 13 and 14). The dashed vertical lines represent the luminances of the two center patches of the dichoptic stimuli. Note that, for calculation of the predictions, the nonlinearity was taken into account; the corresponding values are depicted in the bottom right diagram in Figure 15. See text for further details.

in Figures 13 and 14). Figure 17 shows the respective predictions (red curve) together with the corresponding matching data, averaged across all five subjects (black curve). The two center patch luminances were fixed at 15 cd/m^2 and 35 cd/m^2, respectively. Surround luminances between these values produce inc–dec pairings, stimulus conditions that only produce ON–OFF pairings between eyes (see also Figure 1). Accordingly, perceived and predicted luster magnitudes peaked within this interval of background luminances (see the shaded area in Figure 17). For the inc–inc and dec–dec conditions, which exclusively encompass ON–ON and OFF–OFF pairings, respectively (see the left and right parts of the diagram), the magnitude of perceived luster is considerably lower. This is also predicted by our model, even though the predictions are somewhat higher than the empirical measurements. As expected, the dec–dec configurations generally evoked slightly weaker lustrous impressions compared to dichoptic inc–inc combinations. In order to evoke the same strength of perceived luster, considerably stronger interocular luminance differences would therefore be required for inc–inc or dec–dec stimuli than for inc–dec combinations. Also note the local minima of the prediction line at the two center patch luminances (see the dashed vertical lines in the diagram). In these two cases, one center patch has the same luminance as the surround (and the ring), which means that a center–surround stimulus in one eye is dichoptically combined with a spatially uniform stimulus in the other eye. The model prediction is in agreement with what Kiesow (1920) found in his study. He observed that, under these conditions—that is, with a lack of contrast (and contour) information in one of the monocular half images—no luster occurs (Wendt & Faul, 2022). In order to account for Kiesow's finding, our model requires nonzero LoG filter values at corresponding positions of both images. However, our empirical data seem to contradict this observation in general; corresponding trends can only be found in the data of subject S1 and, to a weaker degree, of subject S5 (see the respective diagrams in Figure 13).

Another important difference between the models is that our former model (implicitly) assumed a linear relationship between perceived luster strength and the amount of interocular conflict. This generally leads to an underestimation of moderate lustrous responses relative to stronger ones. In our present study, the data revealed a strong nonlinear relationship between the two measures, which can be well described by a decelerating power function (Figure 16). Very similar functions were determined in a recent study by Kingdom, Seulami, Jennings, and Georgeson (2019), who measured thresholds for the detection of interocular contrast differences in complex stimuli. The stimuli consisted of horizontally oriented multi-spatial frequency patterns in which local interocular contrast differences were produced by phase shifts between corresponding luminance gratings (Kingdom, Jennings, & Georgeson, 2018). These interocular contrast differences are assumed to be signaled by a lustrous sensation (Wendt & Faul, 2022). The model that Kingdom et al. (2019) proposed to predict their threshold data contains two components: (1) a binocular differencing channel that integrates the two monocular contrast signals, and (2) a nonlinear transducer that is expansive for weaker difference signals and compressive for stronger signals.

In our case, however, the interpretation of these functions as a nonlinear transducer on the interocular conflict signals is problematic, because the dependent variable in the matching experiment (i.e., the interocular luminance or contrast difference) is not a direct measure of perceived luster. Instead, it can be regarded as an (intervening) variable that influences the luster strength in a more or less complex but presumably monotonic way (Wendt & Faul, 2019). A somewhat more appropriate, but also not unproblematic, measure of perceived luster strength is the luster ratings, which were found to strongly correlate with the luster settings (with $r = 0.963$ for the averaged data from all four experiments; see the left diagram in Figure 18). When we look at the relationship between the mean rating data and corresponding conflict measures, the fit with a decelerating power function is even slightly better with $R^2 = 0.829$ (see the right diagram in Figure 18). However, also similar to the study of

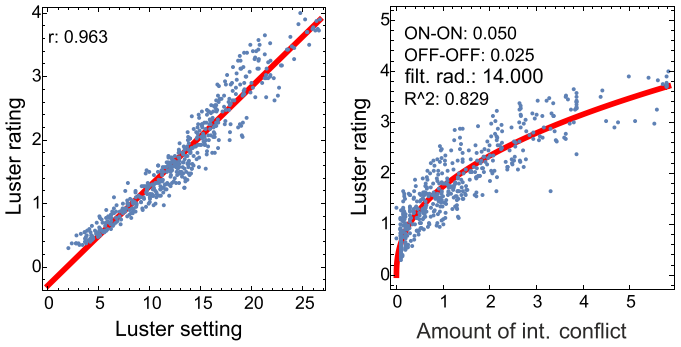

Figure 18. (Left) The mean luster ratings strongly correlate with corresponding luster settings (with $r = 0.963$). (Right) Thus, the relationship between the luster ratings and corresponding conflict values can also be well fitted with a decelerating power function.

Kingdom et al. (2019), we cannot exactly tell at what stage of visual processing this nonlinearity actually occurs. For example, Legge and Foley (1980) suggested very similar nonlinear transducers for monocular contrast signals. It would therefore also be possible that the monocular contrast signals are adjusted by a nonlinear transducer before they are combined into an interocular conflict signal. Moreover, Whittle (1986) found different contrast discrimination functions for incremental and decremental suprathreshold contrasts, suggesting different transducers for the two contrast mechanisms. It would be interesting to investigate whether these different behaviors at the monocular level could explain the different weights for the ON–ON and OFF–OFF pairings in our model.

These different weights can also be interpreted at another level. As mentioned in the Introduction, the Oppel–Helmholtz approach to binocular luster assumes that the visual system interprets the occurrence of a between-eye difference in luminance as caused by specularly reflected light from a glossy surface. Because specularly reflected light always adds to diffusely reflected light, only inc–inc stimuli would be consistent with such a physical causation. The higher weight of the ON–ON pairing could therefore also be explained by a visual mechanism involved in the detection of the material properties of surfaces and specifically tuned to inc–inc combinations for this purpose.

In accordance with Kingdom et al. (2019), we also found that a binocular differencing procedure was sufficient to model our data. However, other studies on the detection of interocular contrast differences have provided evidence for an involvement of a binocular summation channel (Georgeson et al., 2016; Kingdom, Sati, Chang, & Georgeson, 2020). Because a binocular differencing mechanism would generally be more responsive to dichoptic luminance patterns of opposite polarity and a binocular summation mechanism more responsive to same polarity stimulus pairs, one of the reviewers suggested that luster may also result from a balance between these two mechanisms rather than from a differencing process alone.

Potential neural substrates of binocular luster

Our current model is based on the assumption that monocular contrast signals from the two eyes interact in some way at an early binocular stage. Instead of being integrated into a stable binocular contrast signal, two incompatible monocular signals seem to remain in a state of competition, which many researchers believe to be the cause for a lustrous sensation (e.g., Brewster, 1861; Howard, 1995; Sachsenweger, 1960; Sachsenweger, 1961; Wendt & Faul, 2022). This may be due to a binocular mechanism that is unable to combine the conflicting monocular signals.

A quite different explanation of the cause of binocular luster comes from Vladusich, Lucassen, and Cornelissen (2007), who rejected the idea that ON and OFF visual pathways meet at all at a binocular stage, much less at an early level of visual processing. These authors found evidence that the perception of achromatic colors cannot be represented by a single one-dimensional dark–bright continuum but that brightness and darkness (whose physiological bases can be seen in ON and OFF channels, respectively) constitute the dimensions of an internal perceptual two-dimensional space of achromatic colors. According to this higher level approach, stimuli showing reversed contrast polarities between eyes will simultaneously produce percepts of brightness and darkness associated with the same location in the visual field which in their combination may produce the impression of luster.

In contrast to the original version of the neural conflict approach, which is based on the assumption that a conflict results from the inability of the visual system to integrate two discrepant monocular signals, there is some recent evidence from neurophysiology suggesting quite the opposite—namely, the existence of specific binocular cells in early visual cortex that respond strongly to interocular discrepancies and whose purpose is presumably to detect and process interocular conflicts (e.g., Katyal, Engel, He, & He, 2016; Katyal, Vergeer, He, He, & Engel, 2018, who provided evidence that such neurons mediate the alternation process in binocular rivalry). In the case of binocular luster, Kingdom, Read, Hibbard, and May (2022) proposed so-called *tuned-inhibitory cells* as the neural substrate of this phenomenon. Within an ensemble of other binocular cells, this type of neuron seems to play a role in the detection of binocular disparities (e.g., Poggio & Fischer, 1977; Read & Cumming, 2004). In single-unit recordings in the primary visual cortex of primates, the tuned-inhibitory cells were found to strongly respond to anticorrelated random dot stereograms presented with zero disparity—that is, stereograms in

which corresponding dots were black in one eye and white in the other eye and therefore showed reversed contrast polarities between eyes (Zöller, 1998). In contrast, correlated random dot stereograms (stimuli with consistent polarities between eyes) produce only very weak responses in the tuned-inhibitory cells when they are presented with zero disparity (Prince, Cumming & Parker, 2002; Read & Cumming, 2003; Read & Cumming, 2004). Note, however, that in the modeling of the sensitivity functions of these cells, the responsiveness to correlated zero-disparity stimuli was often ignored and set to zero (Cumming & Parker, 1997; Read, 2005; Read, Parker, & Cumming, 2002).

The high responsiveness to luminance features with reversed contrast polarities between eyes is puzzling, though, because, as Read (2005, p. 85) pointed out, "anti-correlated stimuli are unknown in natural viewing." So, why do cells exist in the visual brain that are tuned to impossible stimuli? A possible answer to this question was provided by Read and Cumming (2007), who actually saw a functional purpose in this cell behavior that may contribute to solving the stereo correspondence problem: When searching for corresponding features between the two retinal images, such cells could serve as a kind of "lie detector" whose task within this process would be to detect and exclude false matches. Similarly, Goncalves and Welchman (2017) provided evidence that stereopsis is based on a neural process in which disparity detection is combined with the detection of interocular dissimilarities.

However, assuming that the tuned-inhibitory cells also represent the neural basis of binocular luster, the question remains why the detection of an interocular mismatch is made conscious at all, particularly in the form of a lustrous impression. A possible interpretation is that binocular luster is the "unintended" byproduct of a mechanism whose underlying neural circuitry, although good at detecting binocular disparities under natural conditions, is not designed for artificial stimuli, such as dichoptic stimuli with reversed contrast polarities.

Our present results at least do not contradict the idea that tuned-inhibitory cells are involved in the generation of a lustrous sensation, as suggested by Kingdom and colleagues (2022). Especially the high responsiveness of the cell to stimuli with reversed contrast polarities but a minimal responsiveness to correlated stimuli is in good agreement with the weighting of different polarity pairings in our model. A weight of 1 was chosen for reversed contrast polarity combinations and considerably lower relative weights of 0.025 and 0.05 for the two types of consistent contrast polarity pairings (see the model description). In at least one aspect, though, our model differs from the properties of the tuned-inhibitory cells. In our present model, monocular inputs from cells with circular symmetric receptive fields are used. In contrast, the tuned-inhibitory cells are assumed to receive their inputs from monocular cells with Gabor patch–like ON and OFF receptive fields, in which the antagonistic regions are spatially arranged side by side (e.g., Read & Cumming, 2003; Read & Cumming, 2004; Tsao, Conway, & Livingstone, 2003). We found, however, that the specific type of the filter kernel used does not make much difference, as a modified version of our model that used Gabor filters instead of LoG filters had comparable predictive power as the original version (Appendix B).

Limitations and future work

The output of our present model is the total sum of all local (that is, pixel-by-pixel) conflict values that occur within the target area. Although the use of this *total amount of conflict* measure is uncritical in our present study, because the target areas were nearly constant under all of our experimental conditions, it will produce misleading results when stimuli with differently sized target areas are compared, simply because larger target areas contain more pixels that will contribute to the overall conflict measure. This means that an additional factor would be required in our model that accounts for the size of the target area. One potential solution, as it was already considered in our original version of the model (Wendt & Faul, 2020), would be the use of the *average amount of conflict*, where the total amount of conflict is divided by the number of pixels that represent a non-zero local conflict value. However, because to date little is known about how the stimulus size actually affects the perceived luster strength (Wendt & Faul, 2022), systematic studies are required to examine this dependency in more detail. For example, it was found that the size of the receptive fields of the retinal ganglion cells increases with increasing distance from the foveal area (Cleland, Harding, & Tulunay-Keesey, 1979; Hubel & Wiesel, 1960; Linsenmeier, Frischman, Jakiela, & Enroth-Cugell, 1982), which suggests that the size of the LoG filters in our model has to be adjusted in dependence on the stimulus size. Additionally, it was found that the sensitivity for the detection of interocular contrast differences decreases with increasing retinal eccentricity (Formankiewicz & Mollon, 2009), indicating the use of variable thresholds in our model. All of these aspects will be the subject of future studies by our group.

Furthermore, our model only accounts for lustrous sensations that are caused by stimuli showing interocular differences in luminance; however, there is a lot of evidence that binocular luster can, although to a considerably weaker extent, also be produced by stimuli with a dichoptic difference in chromaticity alone (Dove, 1851; Jennings & Kingdom, 2016; Jung, Moon, Park, & Song, 2013; Kiesow, 1920; Malkoc & Kingdom, 2012; Wendt & Faul, 2019; Yoonessi & Kingdom, 2009; von Helmholtz, 1867; Wendt & Faul, 2022). At present, we

know almost nothing about the physiological basis of these lustrous sensations evoked by color stimuli—at least, there do not seem to exist any color-sensitive retinal ganglion cells (Gegenfurtner, 2003, but see also Brenner & Cornelissen, 1991). Because it has not yet been examined whether the lustrous response to chromatic stimuli also depends on the spatial properties of center–ring–surround configurations in a predictable way, this will also be part of our future investigations.

Conclusions

The finding that particularly strong impressions of luster can be elicited by simple dichoptic center–surround stimuli with reversed contrast polarities between eyes was explained by a mechanism according to which the visual system is unable to binocularly combine corresponding monocular contrast signals with opposite signs (Anstis, 2000). Support for this *neural conflict* idea was provided in a recent study by Wendt and Faul (2020), who developed and tested a model based on such conflicting ON–OFF pairings. However, this approach ignores the fact that weaker lustrous impressions can also result from stimuli that do not produce any interocular ON–OFF pairings. Furthermore, our original model seemed to systematically underestimate lustrous responses evoked by stimuli producing comparatively low magnitudes of interocular conflict.

Here, we have shown that these problems can be largely resolved by two modifications of the original model, First, by the inclusion of ON–ON and OFF–OFF pairings, the model now also well predicts lustrous sensations in stimuli that do not (or not exclusively) produce interocular contrast signals of opposite sign. The relative weights with which these consistent contrast signal pairings contribute to the overall conflict value were found to be significantly lower than the weight of the ON–OFF pairing. Second, by the use of a decelarating nonlinear inducer function, weaker conflict signals are enhanced relative to stronger signals.

The properties of the improved model are also in good agreement with various proposals and findings in related areas. For example, very similar nonlinear response functions have also been found in a study by Kingdom et al. (2019) in which thresholds for the detection of interocular contrast differences were measured. The different weights, as they are used for the three different types of interocular contrast signal pairings, are also supported by some previous studies (Kingdom et al., 2022; Wendt & Faul, 2019). In particular, the recent idea from Kingdom et al. (2022) that the neural basis of binocular luster might be represented by the so-called tuned-inhibitory cells, which strongly respond to ON–OFF pairings but barely to interocular contrast pairings of the same sign, seems to be well compatible with our present results.

Keywords: binocular luster, binocular fusion, contrast

Acknowledgments

The authors thank Alena Rech for her great support in collecting the data and an anonymous reviewer for valuable comments and suggestions on an earlier version of this manuscript.

Funded by a grant (FA 425/3-2) from the Deutsche Forschungsgemeinschaft (German Research Foundation).

Commercial relationships: none.
Corresponding author: Gunnar Wendt.
Email: gunwendt@psychologie.uni-kiel.de.
Address: Institut für Psychologie, Universität Kiel, Kiel, Germany.

References

Anstis, S. M. (2000). Monocular luster from flicker. *Vision Research, 40*, 2551–2556.

Brenner, E., & Cornelissen, F. W. (1991). Spatial interactions in color vision depend on distances between boundaries. *Naturwissenschaften, 78*, 70–73.

Brainard, D. H. (1989). Calibration of a computer controlled color monitor. *Color Research and Application, 14*, 23–34.

Brewster, D. (1861). On binocular luster. *Report of British Association, 2*, 29–31.

Burr, D. C., Ross, J., & Morrone, M. C. (1986). A spatial illusion from motion rivalry. *Perception, 15*, 59–66.

Chua, S. H., Zhang, H., Hammad, M., Zhao, S., Goyal, S., & Singh, K. (2015). ColorBless: Augmenting visual information for colorblind people with binocular luster effect. *Transactions on Computer-Human Interaction, 21*(6), 1–20.

Cleland, B. G., Harding, T. H., & Tulunay-Keesey, U. (1979). Visual resolution and receptive field size: Examination of two kinds of cat retinal ganglion cell. *Science, 205*(7), 1015–1017.

Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature, 389*(18), 280–283.

Dove, H. W. (1851). Ueber die Ursachen des Glanzes und der Irradiation, abgeleitet aus chromatischen Versuchen mit dem Stereoskop. *Poggendorffs Annalen, 83*, 169–183.

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

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nature Reviews Neuroscience, 4*, 563–572.

Georgeson, M. A., Wallis, S. A., Meese, T. S., & Baker, D. H. (2016). Contrast and luster: 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.

Henriksen, S., & Read, J. C. A. (2016). Visual perception: A novel differencing channel in binocular vision. *Current Biology, 26*, R500–R503.

Howard, I. P. (1995). Depth from binocular rivalry without spatial disparity. *Perception, 24*, 67–74.

Hubel, D. H., & Wiesel, T. N. (1960). Receptive fields of optic nerve fibers in the spider monkey. *Journal of Physiology, 154*, 572–580.

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.

Jung, W. S., Moon, Y. G., Park, J. H., & Song, J. K. (2013). Glossiness representation using binocular color difference. *Optics Letters, 38*, 2584–2587.

Katyal, S., Engel, S. A., He, B., & He, S. (2016). Neurons that detect interocular conflict during binocular rivalry with EEG. *Journal of Vision, 16*(3):18, 1–12, https://doi.org/10.1167/16.3.18.

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.

Kiesow, F. (1920). Osservazioni sopra il rapporto tra due oggetti visti separatamente coi due occhi. *Archivio italiano di psicologia generale e del lavoro, 1*, 3–38, 239–290.

Kingdom, F. A. A. (2012). Binocular vision: The eyes add and subtract. *Current Biology, 22*(1), 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.

Kingdom, F. A. A., Read, J. C. A., Hibbard, P. B., & May, K. A. (2022). Special issue: Coding strategies in binocular vision and stereopsis. *Vision Research, 193*, 107989.

Kingdom, F. A. A., Sati, H., Chang, R., & Georgeson, M. A. (2020). Binocular lustre: Does identification of interocular contrast difference require opposite (light-dark) contrasts in the two eyes? *Journal of Vision, 20*(11), 257, https://doi.org/10.1167/jov.20.11.257.

Kingdom, F. A. A., Seulami, N. M., Jennings, B. J., & Georgeson, M. A. (2019). Interocular difference thresholds are mediated by binocular differencing, not summing channels. *Journal of Vision, 19*(14):18, 1–15, https://doi.org/10.1167/19.14.18.

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

Linsenmeier, R. A., Frischman, L. J., Jakiela, H. G., & Enroth-Cugell, C. (1982). Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurements. *Vision Research, 22*, 1173–1183.

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

Marr, D. (1982). *Vision*. New York: W. H. Freeman and Company.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 207*, 187–217.

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

Oppel, J. J. (1854). Ueber die Entstehung des Glanzes bei zweifarbigen, insbesondere bei schwarzen und weißen stereoskopischen Bildern. *Jahresbericht des Frankfurter Vereins, 1854–1855*, 52–55.

Oppel, J. J. (1857). Bemerkungen zur Stereoskopie, ins Besondere zur Erklärung des Glanzes zweifarbiger Bilder. *Poggendorffs Annalen, 100*, 462–466.

Paillé, D., Monot, A., Dumont-Bècle, P., & Kemeny, A. (2001). Luminance binocular disparity for 3D surface simulation. *Proceedings of SPIE, 4299*, 622–633.

Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology, 40*(6), 1392–1405.

Prince, S. J. D., Cumming, B. G., & Parker, A. J. (2002). Range and mechanism of encoding of horizontal disparity in macaque V1. *Journal of Neurophysiology, 87*, 209–221.

Read, J. C. A. (2005). Early computational processing in binocular vision and depth perception. *Progress in Biophysics and Molecular Biology, 87*, 77–108.

Read, J. C. A., & Cumming, B. G. (2003). Testing quantitative models of binocular disparity selectivity in primary visual cortex. *Journal of Neurophysiology, 90*, 2795–2817.

Read, J. C. A., & Cumming, B. G. (2004). Ocular dominance predicts neither strength nor class of disparity selectivity with random-dot stimuli in primate V1. *Journal of Neurophysiology, 91*(3), 1271–1281.

Read, J. C. A., & Cumming, B. G. (2007). Sensors for impossible stimuli may solve the stereo correspondence problem. *Nature Neuroscience, 10*, 1322–1328.

Read, J. C. A., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Visual Neuroscience, 19*, 735–753.

Sachsenweger, R. (1960). Studien über den stereoskopischen Glanz. *Albert von Graefes Archiv für Ophthalmologie, 162*, 518–526.

Sachsenweger, R. (1961). Zum stereoskopischen Glanz. *Albert von Graefes für Ophthalmologie, 164*, 231–234.

Schiller, P. H. (1992). The ON and OFF channels of the visual system. *Trends in Neurosciences, 15*, 86–92.

Sheedy, J. E., & Stocker, E. G. (1984). Surrogate color vision by luster discrimination. *American Journal of Optometry and Physiological Optics, 61*, 499–505.

Tsao, D. Y., Conway, B. R., & Livingstone, M. S. (2003). Receptive fields of disparity-tuned simple cells in macaque V1. *Neuron, 38*, 103–114.

Venkataramanan, K., Gawde, S., Hathibelagal, A. R., & Bharadwaj, S. R. (2021). Binocular fusion enhances the efficiency of spot-the-difference gameplay. *PLoS One, 16(7)*, e0254715.

Vladusich, T., Lucassen, M. P., & Cornelissen, F. W. (2007). Brightness and darkness as perceptual dimensions. *PLoS Computational Biology, 3*(10), 1849–1858.

von Helmholtz, H. (1856). Ueber die Erklärung der stereoskopischen Erscheinung des Glanzes. *Verhandlungen der naturhistorischen Vereinigung der Rheinlande, 13*, 38–40.

von Helmholtz, H. (1867). *Handbuch der physiologischen Optik.* Hamburg: Voss.

Wendt, G., & Faul, F. (2019). Differences in stereoscopic luster evoked by static and dynamic stimuli. *i-Perception, 10*(3), 1–26.

Wendt, G., & Faul, F. (2020). The role of contrast polarities in binocular luster: Low-level and high-level processes. *Vision Research, 176*, 141–155.

Wendt, G., & Faul, F. (2022). Binocular luster – A review. *Vision Research, 194*, 108008.

Whittle, P. (1986). Increments and decrements: Luminance discrimination. *Vision Research, 26*, 1677–1691.

Wienbar, S., & Schwarz, G. (2018). The dynamic receptive fields of retinal ganglion cells. *Progress in Retinal and Eye Research, 67*, 102–117.

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

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.

Zhang, H. (2015). *Spectacularly binocular: Exploiting binocular luster effects for HCI applications* (unpublished doctoral dissertation). National University of Singapore, Singapore.

Zöller, H. (1998). *Stereoskopische Tiefe und stereoskopischer Glanz*. Münster: LIT Verlag.

Appendix A

The improved model requires individual weights for the three different types of interocular contrast polarity combinations: ON–OFF, ON–ON, and OFF–OFF signal pairings. By default, the weight for the ON–OFF pairings was set to 1. The relative weights for the remaining two types of interocular contrast signal pairings were determined in a grid search where we were searching for those combinations of ON–ON and OFF–OFF weights that lead to the best predictions of the empirical luster settings by the conflict measures. Based on previous analyses of our data, the relation between the empirical luster settings and the conflict values was assumed to follow a decelerating power function of the form *luster setting* = a*(*conflict value*)^c. Therefore, we determined the predictive quality of our model in terms of the coefficient of determination R^2 for the best fitting function parameters a and c while systematically varying the weights for the ON–ON and OFF–OFF pairings. Regarding the size of the LoG filter as a further model parameter, it turned out that, at least for those combinations of ON–ON and OFF–OFF weights that resulted in the highest values with $R^2 > 0.817$, the filter radius was quite constant between 13 and 14 pixels. Figure 19 shows the results

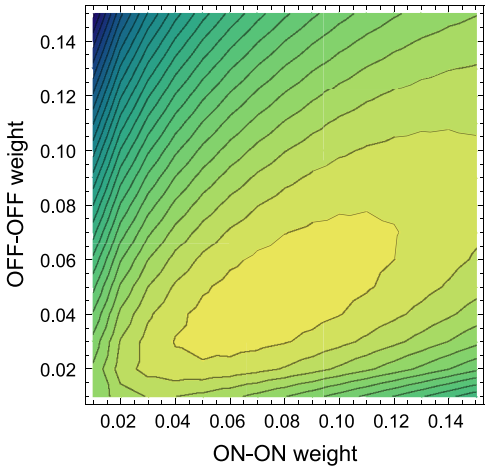

Figure 19. Results of the grid search. In a contour plot, the predictive quality of the model in terms of R^2 is shown in dependence on the ON–ON and OFF–OFF weights (model outputs were calculated with a fixed LoG filter radius of 14 pixels). The contours in the graph represent R^2 values between 0.697 (dark blue region) and 0.817 (yellow region) in steps of 0.005.

of the grid search (calculated with an LoG filter size set to a constant radius of 14 pixels) as a contour plot, in which the R^2 values are shown in dependence on the relative weights of the ON–ON and OFF–OFF pairings. As can be seen, the best predictions (with R^2 values between 0.817 and 0.820 in the yellow area in the contour plot) were obtained with weights for the ON–ON signals within a range between about 0.04 and 0.12 and weights for the OFF–OFF signals that were about half as large as those for the ON–ON pairings.

As there is a comparatively large set of combinations of the two different weights that would be suitable to be used in our model, we also compared the point clouds of our data, in which the empirical luster settings are plotted against corresponding conflict values. In the left diagram in Figure 20, the luster settings are shown in dependence on the amount of interocular conflict as they would result from our former model, where only ON–OFF signals were taken into account (with the ON–ON and OFF–OFF weights set to 0). One can see a large set of data points lined up at a conflict value of 0. This means that there is a large range of empirical luster judgments for which our former model wrongly predicts no luster at all. Also, the predictive quality of this model is comparatively poor (when the data are fitted with a power function) with an R^2 value of 0.594 (calculated with an LoG filter radius of 16 pixels, which provided the highest R^2 value in this case). Such wrong predictions are exactly what we wanted to overcome by including ON–ON and OFF–OFF signal pairings in our model. The middle diagram in Figure 20 shows the same empirical luster settings but now in dependence on conflict values in which the ON–ON and OFF–OFF signals were taken into account with weights of 0.05 and 0.025, respectively. One can see that, with this addition to the model, the wrong "no luster" predictions disappeared and that the data points now conform comparatively closely to the prediction line (red curve). The predictive quality under

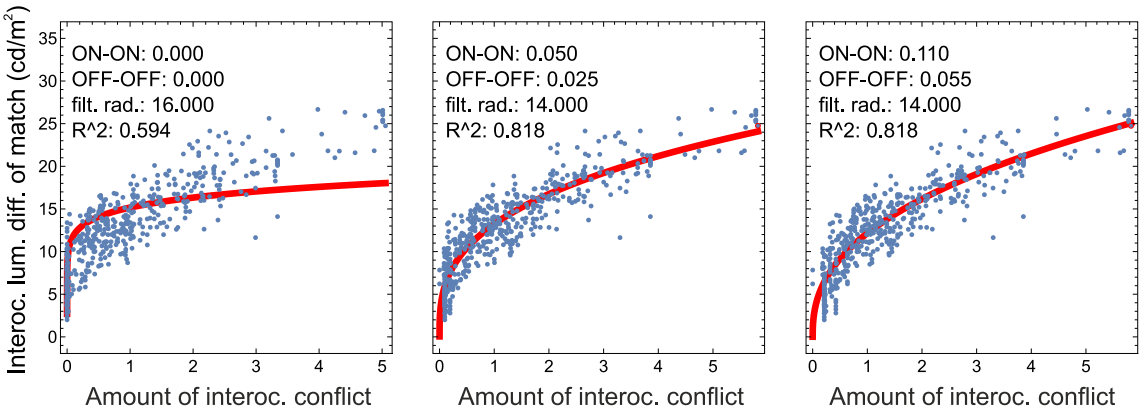

Figure 20. The empirical luster settings of all of our four experiments are plotted against corresponding conflict values separately for different weights for the ON–ON and OFF–OFF signal pairings. In the left diagram, these weights were both set to 0, which means that exclusively ON–OFF pairings were taken into account for the calculation of the model predictions (which is equivalent to the procedure used in our original model; see Wendt & Faul, 2020). Under these conditions, there is a large range of luster settings for which the model wrongly predicts "no luster" (see the array of data points at the conflict value 0). Also the fit with a power function is comparatively poor. Considerably better fits are obtained with conflict values in which ON–ON and OFF–OFF signals are also taken into account (middle and right diagrams). The weights used for these signal pairings particularly affect the position of the data points at the lower end of the conflict scale—that is, at weaker amounts of interocular conflict, for which the visual system seems to be most sensitive.

these conditions has also considerably improved (R^2 = 0.818). The right diagram in Figure 20 shows the luster settings in dependence on the conflict values that were calculated using higher weights for the ON–ON and OFF–OFF pairings, which were 0.11 and 0.055, respectively. Overall, the predictive quality was as high as with the lower weights (R^2 = 0.818); however, the data points at the lower end of the conflict measures are markedly shifted away from the origin. Because the visual system seems to be particularly sensitive to lower amounts of interocular difference (e.g., Kingdom et al., 2019; Wendt & Faul, 2019), a finer resolution of the conflict measure is required here. We therefore decided to use the lower weight combination for the ON–ON and OFF–OFF signal pairings as fixed parameters in our model (see the model description section).

Appendix B

As an alternative to ON- and OFF-center cells with a circular symmetric structure, on which our current model is based, we tested whether our model would also work with monocular signals produced by ON and OFF cells whose receptive fields have Gabor patch–like, side-by-side, spatially antagonistic regions and therefore are orientationally selective. Such kind of cells are assumed to provide monocular signals that are integrated in the binocular tuned-inhibitory cells (Read & Cumming, 2003; Read & Cumming, 2004; Tsao et al., 2003).

To this end, the receptive fields of these monocular cells were simulated using a Gabor filter with an elongated OFF-center area flanked by ON surround areas at both sides (see the inlets between steps 1 and 2 in Figure 21). We used the Mathematica function GaborFilter with filter radius r; a standard deviation of the Gaussian envelope of sigma = $r/2$; wave vectors 1.5 PI/r (1,0) and 1.5 PI/r (0,1) for the horizontally and the vertically oriented kernel, respectively; and a phase shift of PI. This filter kernel was adjusted such that its weights added up to zero. Analogous to the logic used in the LoG-based model (see the model description), a stimulated ON-center cell would be represented by a negative filter value and an OFF-center cell by a positive value. Following essentially the same processing steps as in the LoG version of our model, the two half-images of a dichoptic stimulus pair (steps 1 to 2 in Figure 21) were first convolved with the OFF-center Gabor filter. Because this filter kernel is orientationally selective, we separately convolved each half-image (which all have a square-shaped, center–ring–surround structure) with a horizontally and a vertically oriented filter (step 2 in Figure 21). In order to identify ON–OFF, ON–ON,

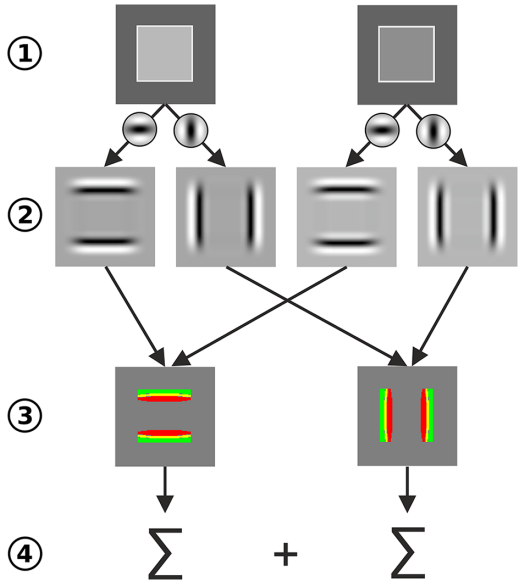

Figure 21. Schematic representation of how, in a modification of our model, the total amount of conflict measure was determined using Gabor filters instead of LoG filters. Each of the two half-images of the dichoptic stimulus (1) was first separately convolved with a horizontally and a vertically oriented Gabor OFF-center filter kernel (see the inlets between 1 and 2). Corresponding component images were then analyzed to identify either ON–OFF (yellow parts in 3), ON–ON (green parts), or OFF–OFF pairings (red parts) between them on a pixel-by-pixel basis. The single pixel-based conflict values were then summed up for each of the two filter orientation components, taking into account the different weights as described in Appendix A and the model description section. As a measure for the total amount of conflict induced by the entire stimulus, the sum of these two components was used (4).

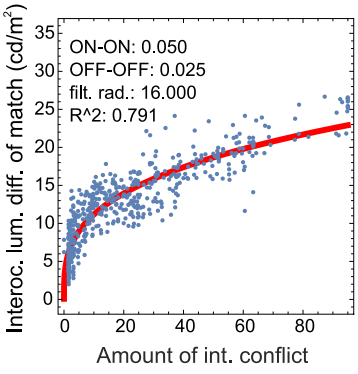

Figure 22. The mean luster settings of all four experiments are plotted against the corresponding conflict values as they were calculated by the model using Gabor filters instead of LoG filters (cf. Figure 19). The red curve shows the fit with a power function.

or OFF–OFF pairings, corresponding pixels within the target area of the half-images were analyzed with respect to the sign of their filter values, separately for the two different orientation component image pairs (step 3 in Figure 21). The resulting total amounts of conflict measures for the two orientation components were then added to obtain a conflict measure representing the entire stimulus pair (step 4 in Figure 21).

The diagram in Figure 22 shows in the usual way the luster settings from all four experiments, averaged across all five subjects, in dependence on the corresponding interocular conflict measure as they were obtained with the Gabor filter version of our model. Again, the relationship between the two measures can be well fitted with a decelerating power function. The proportion of explained variance of about 79.1% (obtained with a filter radius of 16 pixels) is comparable to that obtained with our LoG-based model ($R^2 = 0.818$), suggesting that the underlying monocular mechanisms can indeed be appropriately modeled with receptive fields based on orientation-selective Gabor patches.