A Bayesian observer model reveals a prior for natural daylights in hue perception

Yannan Su[1,2], Zhuanghua Shi[3], and Thomas Wachtler[1,4]

[1]Faculty of Biology, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

[2]Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

[3]General and Experimental Psychology, Ludwig-Maximilians-Universität München, Munich, Germany

[4]Bernstein Center for Computational Neuroscience Munich, Planegg-Martinsried, Germany

Feb. 2023

Abstract

Incorporating statistical characteristics of stimuli in perceptual processing can be highly beneficial for reliable estimation to overcome noisy sensory measurements but may generate perceptual bias. According to Bayesian inference, perceptual biases arise from integrating non-uniform internal priors with noisy sensory inputs. We used a Bayesian observer model to derive biases and priors in hue perception based on discrimination data for hue ensembles with varying levels of chromatic noise. For isoluminant stimuli with hue defined by azimuth angle in cone-opponent color space, discrimination

thresholds showed a bimodal pattern, with lowest thresholds near a non-cardinal blue-yellow axis that aligns closely with the variation of natural daylights. Perceptual biases showed zero crossings around this axis, indicating repulsion away from yellow and attraction towards blue. The results could be explained by the Bayesian observer model through a non-uniform prior with a preference for blue. Our results indicate that observers exploit knowledge of colors in natural environments in visual processing for hue perception.

Author Summary

Human visual perception is susceptible to systematic biases consistent with an optimal inference process that combines sensory evidence with prior knowledge. It has been hypothesized that, to make accurate inferences, the brain adapts to the sensory world and develops a prior that reflects environmental statistics. Our research on hue perception supports this idea, which demonstrates that color discrimination exhibits a consistent preference towards blue and away from yellow, which can be attributed to the influence of natural daylights that are dominated by blue and yellow. Our results are predicted by a Bayesian model incorporating prior knowledge about natural daylights, which provides insight into how humans adapt to these lighting conditions and use natural color statistics in hue perception.

1 Introduction

Acquiring and processing complex sensory information can be a time and energy-intensive task. However, we are fortunate enough to live in a world that exhibits regularities, which can aid us in saving energy and processing resources. Prior knowledge gained through recognizing patterns and statistical dependencies can greatly enhance our ability to process incoming sensory information. For instance, bananas are known to appear yellow depending on their ripeness. Hence, if we consider the task of recognizing bananas, it would be based on the color yellow rather than, say, purple.

Using previously acquired knowledge during perceptual processing can be traced back to Helmholtz's idea of 'unconscious inference' (von Helmholtz, 1867). In recent decades, the development of the Bayesian inference framework suggests that incorporating prior knowledge can significantly enhance the reliability of perceptual estimation, especially when the input signals are corrupted by noise (Knill and Pouget, 2004; Shi et al., 2013). Within the Bayesian framework, an observer's statistical inference is influenced by two key components: the likelihood and the prior (Knill and Pouget, 2004), which are optimally integrated according to their reliability. While prior knowledge can be valuable in many situations, it can also create perceptual biases. Perceptual biases occur when the prior mismatches the likelihood.

A common difficulty encountered in the Bayesian inference framework lies in measuring likelihood and prior, which are not directly accessible. Stocker and Simoncelli (2006) proposed a method to recover likelihood and prior from psychophysical measurements of perceptual bias and variability. These measurements were assumed to reflect the location and spread of the estimated posterior distribution. Specifically, Stocker and Simoncelli (2006) obtained these measurements from discrimination between stimuli with different likelihood widths, enabling observing discrimination bias and deriving the underlying prior. Girshick et al. (2011) further applied this approach to investigate orientation perception and revealed that a perceptual prior of orientations peaked at the cardinal orientations that are the dominant orientations in natural scenes, suggesting that visual perception may involve using prior information regarding the regularities of natural scene structures.

The Bayesian inference framework has successfully accounted for many perceptual biases in object perception (Kersten et al., 2004), multi-sensory integration (Ernst and Banks, 2002), sensorimotor

learning (Körding and Wolpert, 2004), visual speed perception (Stocker and Simoncelli, 2006), visual orientation perception (Girshick et al., 2011), and time perception (Shi and Burr, 2016; Glasauer and Shi, 2022). However, it has rarely been used to investigate color vision. Attempts focusing on color constancy (Brainard and Freeman, 1997; Brainard et al., 2006) had difficulties determining adequate priors to explain the experimental findings. Various studies examined *variability* in hue perception (Boynton et al., 1986; Krauskopf and Gegenfurtner, 1992; Danilova and Mollon, 2010; Witzel and Gegenfurtner, 2013) or *biases* in hue perception induced by spatial or temporal context (Olkkonen et al., 2014; Klauke and Wachtler, 2015; Kellner and Wachtler, 2016). The existence of internal biases in hue perception remains unknown. However, there is evidence that color perception is adapted to the statistics of the natural environment (Macleod, 2003; Mollon, 2006; Bosten et al., 2015; Klauke and Wachtler, 2015), and similar adaptation is thought to underlie in orientation vision perceptual biases towards the cardinal orientations in natural scenes (Girshick et al., 2011). If common principles of neural processing underlie perceptual biases observed in different visual domains (Klauke and Wachtler, 2015), it could be expected that color vision exhibits internal biases reflecting regularities in natural colors. To understand how prior knowledge of natural daylights is integrated into color processing, we therefore measured perceptual variability and biases through discrimination of noisy hue ensembles by different individuals and applied the Bayesian observer model to recover their priors.

2 Results

2.1 Behavioral Results

We used two-alternative forced-choice discrimination experiments to measure perceptual variability and bias in hue perception. Observers were presented with two arrays of noisy hue patches and asked to compare their average colors (Fig. 1a). They had to indicate whether the ensemble averages of the arrays from left to right matched the direction of color change shown by a gradient color bar displayed above the stimuli (see more details in the section Methods). A stimulus pair consisted of a reference array (with a fixed average hue across trials) and a comparison array (with an average

hue adjusted by a staircase method). For each of the eight reference stimuli (Fig. 1b), we tested three conditions: two same-noise conditions (low-noise versus low-noise, L-L, and high-noise versus high-noise, H-H), and one cross-noise condition (low-noise versus high-noise, L-H).

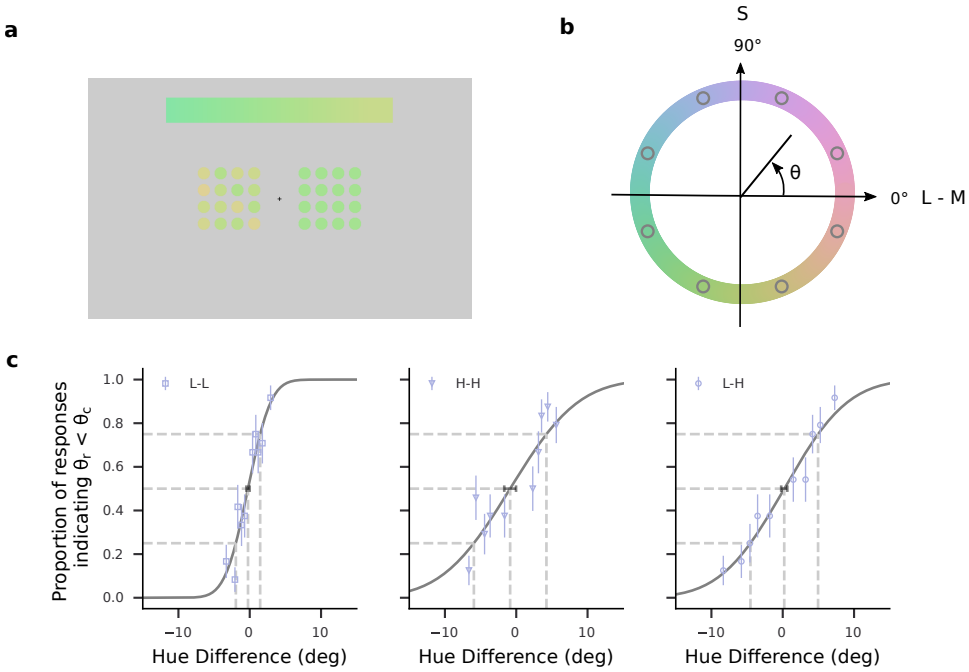

Figure 1: Experiment and psychometric functions. (a) Example of a stimulus display. Participants were asked to compare two arrays of color patches, presented on the left and right side of fixation, and to indicate whether their respective average hues matched the direction of hue changes that was indicated by the color gradient bar at the top of the display. The figure shows the cross-noise condition with a high-noise array on the left and a low-noise array on the right. Note that hue differences have been exaggerated in the figure for illustrative purposes. (b) Eight reference hues were defined by azimuth angle θ in opponent cone-contrast color space (gray circles). (c) Examples of psychometric functions for the three noise conditions. Proportions of responses indicating the subject perceived the hue angle of the comparison stimulus (θ_c) as larger than the hue angle of the reference stimulus (θ_r) are plotted as a function of the difference between the hue angles θ_c and θ_r. Data are from a single subject's responses with $\theta_r = 112.5°$ for the three noise conditions L-L (left), H-H (center), and L-H (right). Error bars of the data points denote standard error. Solid lines show cumulative Gaussian functions fitted to the data. Dashed lines denote 25%, 50%, and 75% values. Horizontal error bars around the estimated 50% points denote 68% confidence intervals.

We then fitted the psychometric data for hue discrimination across varying noise conditions with cumulative Gaussian functions (Fig. 1c). Based on the estimated psychometric function, we used the standard deviation of the cumulative Gaussian function as the just noticeable difference (JND)

and calculated the point of subjective equality (PSE) at the threshold of 50%. These values reflect the discrimination thresholds and bias of each participant, respectively.

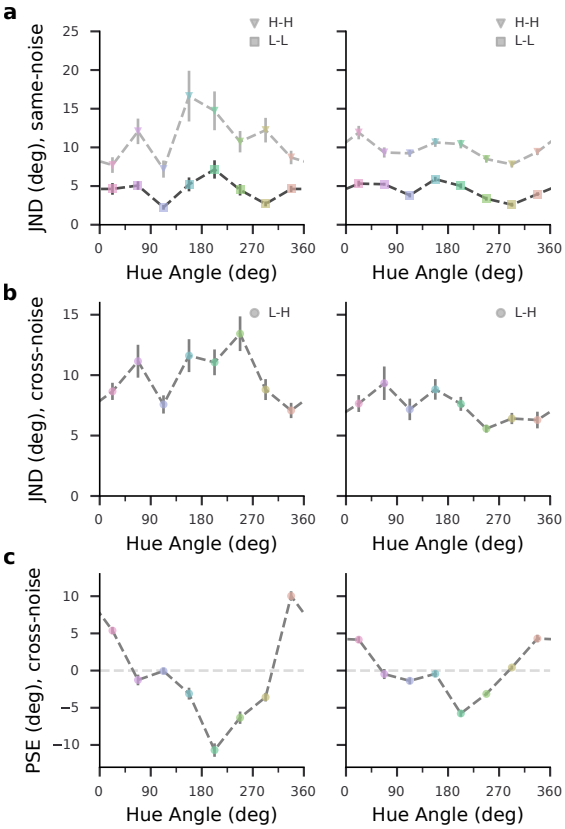

Figure 2: Experimental data for a single subject (left) and the average subject (right). (a) Hue discrimination thresholds (JNDs) under the same-noise condition. (b) Hue discrimination thresholds under the cross-noise condition. (c) Biases under the cross-noise condition, measured as hue angle differences between the high-noise and low-noise stimuli at the PSE. Bars denote one standard error of the estimates.

Discrimination thresholds

Under the same-noise conditions, discrimination thresholds as a function of hue angle typically exhibited a bimodal pattern (Fig. 2a). On average, across all subjects, the L-L condition had two local minima at the hue angles, $101.3^\circ \pm 22.5^\circ$ and $298.1^\circ \pm 28.8^\circ$ (see Fig. S1 and S2), and two local maxima occurred at $49.5^\circ \pm 24.6^\circ$ and $170.4^\circ \pm 21.9^\circ$. Fitting a sine mixture model to the discrimination thresholds yielded similar hue angles corresponding to these extrema (local minima

at 97.3°±28.4° and 290.7°±18.1°, and local maxima at 39.4°±47.1° and 186.7°±23.6°, averaged across subjects; Figure 3a-b, also see Fig. S4a). Thus, discrimination thresholds followed a bimodal pattern, with the lowest thresholds near an axis connecting blue (112.5°) and yellow (292.5°) and with the highest thresholds at the axis orthogonal to the blue-yellow axis. This is consistent with results of previous studies (Danilova and Mollon, 2010; Witzel and Gegenfurtner, 2013). The bimodal pattern persisted in the H-H condition, but of the discrimination thresholds were significantly higher (sign test $p < .001$) in the H-H condition than in the L-L condition. The cross-noise condition (L-H) had intermediate discrimination thresholds (Fig. 2b). Of the L-H discrimination thresholds, 82.5% were higher (sign test $p < .001$) than the corresponding L-L discrimination thresholds, and 90.0% were lower (sign test $p < .001$) than the H-H discrimination thresholds.

Biases

We observed non-negligible bias only under the cross-noise condition (L-H) for all subjects (Fig. 2c; also see Fig. S3 for the same-noise bias). Given that in the cross-noise condition, the low-noise stimulus always served as the reference stimulus, with fixed hue across trials, the cross-noise bias shown in Fig. 2c represents the perceptual bias of the high-noise stimulus relative to the low-noise stimulus. Averaged across subjects, the biases showed a minimum of $-5.6° \pm 3.1°$ and a maximum of $5.5° \pm 2.9°$. Among the biases, 88.9% of the negative values corresponded to hue angles within the range of 112.5° to 292.5°, and 85.7% of the positive values were at the hue angles smaller than or equal to 112.5°, or larger than or equal to 292.5°. Subjects' biases showed zero-crossings at the hue angles within 45° around 112.5° and within 45° around 292.5°, except for one subject whose zero-crossings did not fall within the hue angles of $112.5° \pm 45°$. Around the zero-crossings biases were attractive towards blue hues (near 112.5°) and repulsive away from yellow hues (near 292.5°). Most subjects exhibited only two zero-crossings around 112.5° and 292.5°, while two subjects showed additional zero-crossings around 157.5° (see S2 and S6 in Fig. S1). According to the Bayesian observer model, perceptual bias arises when the prior is non-uniform and misaligned with sensory measurements. Moreover, assuming that the same prior is used, if the width of the likelihood is different, the posterior will be different. This prediction matches our observations of non-zero cross-noise bias, suggesting non-uniformity of the perceptual prior.

Nevertheless, it is possible that the cross-noise bias may result from non-uniform discrimination thresholds. If changes in sensitivity occur over a scale of hue angles comparable to the hue range of the noisy stimulus arrays, then the perceived ensemble average for a given stimulus may vary depending on the noise level. Thus, the cross-noise bias might reflect not only the prior but also the local non-uniformity of discrimination. Therefore, different biases could arise depending on whether the comparison stimulus had a larger or smaller hue angle value than the reference stimulus. To rule out this possibility, we simulated the effect of non-uniform thresholds for the data of a representative subject S3. We split the data from the L-L condition into two groups based on whether the comparison stimulus had a larger or smaller hue angle value than the reference stimulus. We computed the ratio of 75%-correct thresholds between the two groups and multiplied the stimulus hue angle by the inverse of this ratio. Thus, we scaled the comparison stimuli in the cross-noise condition and estimated the cross-noise bias by fitting a psychometric function to the scaled data. We found that the cross-noise bias was unchanged by scaling (sign test $p = .46$), which rejects the possibility that non-uniform discrimination thresholds account for the cross-noise bias.

Taken together, we found the sensitivity of hue discrimination followed a bimodal pattern and observed a systematic bias for perceived color, with observers showing the best discrimination and the least bias for bluish and yellowish stimuli. The introduction of chromatic noise resulted in an increase in discrimination thresholds and a modulation of the bias. Moreover, the results also suggest that the observed bias may be attributed to non-uniform internal priors associated with the axis connecting blue and yellow in color space (Mollon, 2006).

2.2 Modeling Results

To determine priors employed by observers in color hue discrimination, we used a Bayesian ideal observer model and optimized the parameters of its prior to account for the data from our human observers.

The model connects two behavioral measurements (discrimination threshold and bias) and two Bayesian components (likelihood and prior). Specifically, the stimulus uncertainty was propagated from the measurement distribution to the posterior distribution, resulting in perceptual variability

(Girshick et al., 2011). In line with this model, measurement distributions and likelihood functions were computed from the fitted same-noise variabilities. When stimulus noise increased the threshold, the widths of the corresponding measurement distribution and likelihood function were also increased (Fig. 3, see also Fig. S4). We assumed the ideal observer's estimates of a particular stimulus θ corresponded to maximum a posteriori (MAP) estimates $\hat{\theta}$ resulting from multiplying the likelihood function with a prior at the Bayesian decoding stage. We simulated each subject's cross-noise data by comparing each pair of MAP estimate distributions ($\hat{\theta}_L$ and $\hat{\theta}_H$). The prior was modeled as a von Mises function, and the prior parameters were obtained by maximizing the likelihood of the experimental data given the simulation-generated psychometric function. For the average data, the estimated prior showed the highest values at $115.2° \pm 19.9°$ (Fig. 4). The cross-subject average of the hue angle where the prior peaked was $107.3° \pm 26.7°$ (SE) (see also Fig. S5), that is, close to unique blue.

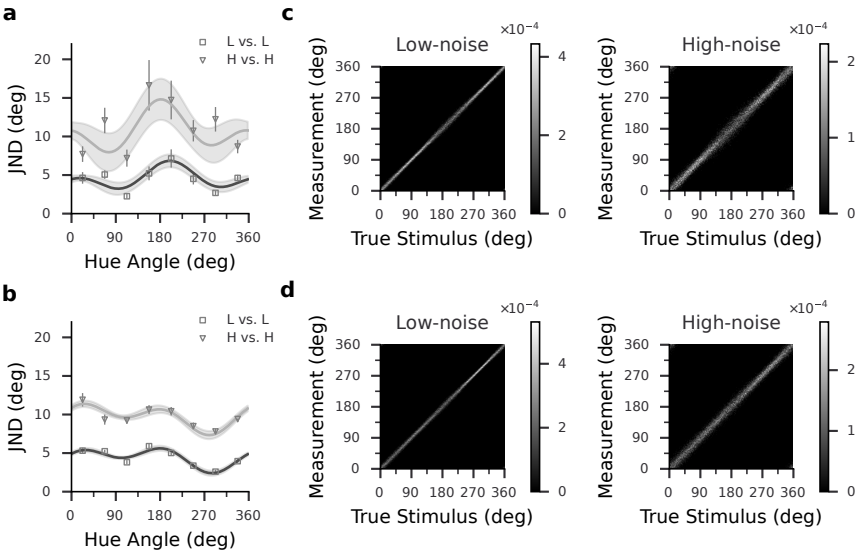

Figure 3: Threshold fits and estimated likelihood functions in the same-noise conditions. (a-b) Fitted JNDs of subject S3 (a) and the average subject (b). JND estimates are from the data shown in Fig. 2, error bars indicate one standard error. The dark and light gray lines are the fitted JNDs for the L-L and the H-H conditions, respectively. The gray shaded area indicates 68% confidence intervals of fitted JNDs. (c-d) Estimated likelihood functions of subject S3 (c) and the average subject (d). Each horizontal slice of the two-dimensional function represents a likelihood function of stimulus hue angle θ given a particular measurement m, and each vertical slice represents a measurement distribution centered on a particular θ. The gray level represents corresponding probability densities.

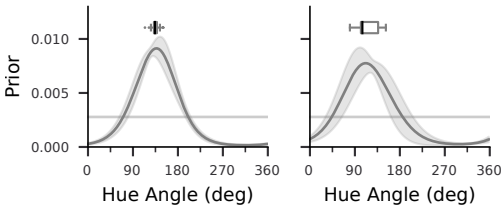

Figure 4: Estimated prior for subject S3 (left) and the average subject (right). The gray shaded area indicates ±1 one standard deviation of 100 bootstrapped estimates. The boxes above the curves indicate the first quartile to the third quartile of the peak locations of the 100 bootstrapped estimates, with the black line at the median. The whiskers extend from the box by 1.5× the interquartile range (IQR). Flier points indicate values beyond the range of the whiskers. The light gray horizontal line represents the uniform prior.

To assess the effectiveness of the encoding-decoding model with the estimated prior, we compared its performance in predicting the cross-noise bias against a model with a uniform prior. The model with the estimated prior was found to be better at predicting both the sign and amplitude of the cross-noise bias than the model with the uniform prior (Figure 5a-b). Across all subjects, the normalized log-likelihoods indicated that the estimated prior outperformed the uniform prior (Fig. 5c). Furthermore, when comparing the two prior models using Akaike Information Criterion (AIC) scores (Akaike, 1998), we found that the estimated prior consistently performed better than the uniform prior (Supplementary Table S1). Additionally, modeling the prior as a normalized sine function yielded similar results, with the estimated prior peaking around the blue hue ($119.1° \pm 28.6°$ (SE), averaged across subjects, see Fig. S7). Note that the sine function had a constrained period of 360° and thus exhibited a single peak over the hues. Its prediction performance was better than a bimodal sine function with period 180° (Fig. S7), which confirmed the unimodality of the prior.

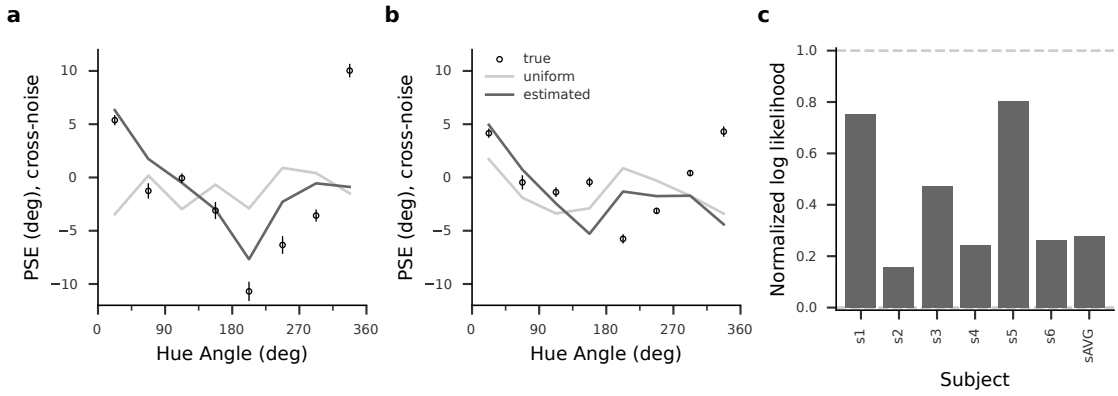

Figure 5: Comparison of priors. (a-b) Cross-noise biases with predictions from estimated prior (black lines) and uniform prior (gray lines) for subject S3 (a) and the average subject (b). Circles represent the cross-noise biases shown in Fig. 2. (c) Normalized log-likelihoods of predictions using the estimated prior for all subjects, including the average subject (sAVG). Values greater than 1 indicate prediction better than raw psychometric fits, and values greater than 0 indicate prediction better than the model with uniform priors.

In summary, these results indicate that the observers used a non-uniform prior related to the blue-yellow axis. Specifically, the bell-shaped prior showed the highest probability at blue and the lowest probability at yellow.

3 Discussion

Our study aimed to investigate the perceptual characteristics of hue discrimination and to identify an internal prior that contributes to hue perception. The results showed that the lowest discrimination thresholds and smallest biases occurred for stimuli near a perceptual blue-yellow hue axis. In addition, we found a cross-noise perceptual bias that was attractive towards blue and repulsive from yellow, which was predicted by a Bayesian observer model with a prior that favored blue. Our study extends the Bayesian perspective on perception to the domain of color and provides evidence of a systematic bias in color perception related to natural daylights.

Ensemble hue perception

Our psychophysical results showed that subjects could effectively integrate the information over noisy hue ensembles, which agrees with previous findings on ensemble hue perception (Maule et al., 2014; Webster et al., 2014; Maule and Franklin, 2015; Virtanen et al., 2020). Specifically, our data confirm that hue averaging does not require a spatial configuration with abutting hue elements (Virtanen et al., 2020).

Previous studies have indicated that categorical effects may influence the percept of hue ensembles (Maule et al., 2014). However, discrimination thresholds in our experiments were approximately equal between yellow and blue ($p = .35$), despite their categories spanning different hue angle ranges (Webster et al., 2000b; Hansen et al., 2007; Witzel and Gegenfurtner, 2013). Moreover, observers were instructed to discriminate hue using an external physical reference instead of internal criteria, minimizing any potential categorical effects from subjective color naming. Thus, any biases we observed here cannot be attributed to categorical effects.

Relation between perceptual variability and bias

The two primary psychophysical measures - perceptual variability and bias - covaried in our experiments: both discrimination thresholds and biases were lowest near the daylight blue-yellow axis and largest near the axis orthogonal to the blue-yellow axis. This finding is consistent with previous studies on orientation perception, which have shown a similar relation between threshold and bias minima occurring at cardinal orientations (Tomassini et al., 2010; Girshick et al., 2011).

However, as previous studies have reported, orientation stimuli with high discrimination thresholds or high variability, such as oblique orientations, can also be perceived with minimal bias (Tomassini et al., 2010; Girshick et al., 2011). Wei and Stocker (2017) presented a mathematical description of the relation between variability and bias, suggesting a proportional relationship between bias and the slope of the square of the discrimination threshold, and found that the relation holds for many visual features, including orientation, motion direction, magnitude, and spatial frequency (Wei and Stocker, 2017). Derived from this model, bias is minimal at all extrema of discrimination thresholds, while attraction bias occurs at the maxima and repulsion bias occurs at

the minima of discrimination thresholds. According to this prediction, one would expect that the bias of perceived hues in our study would show four zero-crossings, corresponding to the number of maxima and minima in the bimodal pattern of discrimination thresholds. Specifically, the hue percept should be biased away from the blue-yellow axis, where discrimination thresholds are minimal, and attracted towards colors with high discrimination thresholds, such as reddish and greenish hues. However, these predictions did not match the measured cross-noise biases well (Fig. S8).

Potentially, our measurements might have missed to consistently identify some zero crossings, particularly around 157.5° where inter-subject variability occurred in biases (Fig. 2c) and some subjects showed more than two zero-crossings (Fig. S1). However, the bias around blue in our results is inconsistent with the relation. Could this indicate that the perception of color is governed by fundamentally different principles than other visual features? Given that so far empirical findings rather point to striking similarities in processing between color and other visual features (Klauke and Wachtler, 2015), we consider this possibility quite unlikely. An alternative possibility that could explain this difference is a strongly skewed likelihood function. Intuitively, one would expect the bias to be zero when the prior of a Bayesian observer is uniform. However, the model with a uniform prior predicted non-negligible bias for some subjects (Fig. 5, see also S2 and S6 in Fig. S6). These results are likely related to the asymmetry of the likelihood function in the observer model (Wei and Stocker, 2015), which resulted from estimating the likelihood directly from the experimental data by sampling from a measurement distribution. The heavy tail of the likelihood might lead to a deviation of the posterior from the true stimulus, such that both likelihood shape and prior could contribute to perceptual bias (Stocker and Simoncelli, 2006; Wei and Stocker, 2015). Nevertheless, the key factor in yielding the systematic bias in our experiments was the non-uniformity of the prior, as the asymmetric likelihood with a uniform prior generates less accurate predictions than our model.

Inter-subject variability

While the bias pattern and prior model were consistent for most subjects, we observed individual differences in the pattern of the cross-noise bias and, consequently, in the peak locations of the

estimated priors. It is thus necessary to consider how well the average data can represent the data from individuals. We determined the results for a hypothetical average subject by pooling the data from all subjects. This average subject showed an average performance: there was no significant difference between the average subject's psychophysical estimates and the mean values of all subjects' psychophysical measurements (sign test $p = 0.44$ for discrimination thresholds, sign test $p = 0.68$ for biases).

Given that considerable individual differences are found in color vision (Webster et al., 2000a, 2002; Witzel and Gegenfurtner, 2013; Emery et al., 2017), one might argue that the variations of the recovered prior's peak location reflect inter-subject variations of color perception. To test whether the location of the prior covaries with other color features, we compared the locations of the prior and color categories of the individual subjects. We determined the unique hue locations in five subjects (see Supplementary Methods) and found the standard deviation in unique hue locations ($5.31° \pm 2.05°$, averaged across four unique hues) was much smaller than that in the prior peak locations (70.46°). For one observer (S1) whose unique blue and yellow hue locations deviated from other subjects' unique hues and shifted towards larger hue angles, the corresponding prior peaked at a hue angle close to the average peak location of all subjects' priors. Moreover, another observer (S6) with the prior peaking at the largest hue angle among subjects did not show such pronounced deviations in unique hue categories. We did not find a correlation between the prior peak location and any of the unique hue locations (Pearson correlation $r = -0.33, p = 0.59$ for unique blue, $r = -0.77, p = 0.12$ for unique yellow, $r = -0.09, p = 0.88$ for unique red, $r = -0.42, p = 0.47$ for unique green). Thus, we do not have strong evidence supporting that individual variation in unique hues accounts for the inter-subject variations of priors.

Alternatively, we consider that the variations might result from the inter-subject variation in the color statistics of the natural environment, which was propagated into a prior learned through subjective perceptual experiences within the Bayesian framework. From the perspective of evolution, color perception may adapt to the regularities of colors in the natural world (Shepard, 1992). Even for the same observer, the long-term adaptation to different natural scenes could potentially result in variations in color appearance (Juricevic and Webster, 2009). Thus, the inter-subject variation we observed will likely be formed over long timescales through subjective perceptual experience.

Hue perceptual non-uniformity and blue-yellow asymmetries

Perceptual quantities as measured in our study are not uniform over the physiological color space based on cone-opponent coding. Specifically, discriminability is high, and variability is low around the blue-yellow axis (Boynton et al., 1986; Krauskopf and Gegenfurtner, 1992; Danilova and Mollon, 2010; Witzel and Gegenfurtner, 2013; Klauke and Wachtler, 2015). Our results verified that such non-uniform discriminability also exists for hue ensembles with chromatic noise. More importantly, we found that hue judgments were least biased near blue and yellow, further reflecting the non-uniformity. The non-uniformity arises not along the 'cardinal' axes of precortical cone-opponency, but an oblique axis connecting perceptual unique blue and yellow. This axis coincides with the variation of natural daylight (Mollon, 2006) with corresponding high variation in cone signals (Webster and Mollon, 1997; Wachtler et al., 2001). It has a pronounced representation in visual cortex (Wachtler et al., 2003) and has specific roles in color induction (Klauke and Wachtler, 2015) and color constancy (Pearce et al., 2014).

While blue and yellow show symmetries, in terms of their similarities in perceptual quantities and the coincidence with the daylight locus (Mollon, 2006; Webster, 2020), the bias in our results was attractive towards blue and repulsive from yellow, indicating asymmetries between blue and yellow. A previous study has demonstrated a systematic deviation towards blue when subjects adjusted yoked blue-green hue pairs to achieve an equal perceived mixture of binary hues (Webster et al., 2014). The deviation occurred for unique blue settings and not for other unique hues, which matches the unimodality of the systematic prior revealed in our results. Moreover, the deviation only occurred in the blue-green settings, while no conspicuous bias arose in the mixture consisting of yellow hues, which strongly evidenced asymmetries between blue and yellow.

A special role of blue has been found previously in color constancy: bluish illumination results in higher color constancy than other chromatic illuminations. The so-called "blue bias" has been evidenced by the observations that slightly bluish illuminants tend to be perceived as neutral (Delahunt and Brainard, 2004; Pearce et al., 2014; Winkler et al., 2015; Radonjic et al., 2016; Aston et al., 2019), and achromatic adjustments are shifted toward blue (Weiss et al., 2017). An explanation for the blue bias was that the illumination sensitivity threshold is higher for blue than for other

colors (Pearce et al., 2014; Radonjic et al., 2016), which would be in line with the hypothesis that the visual system may adapt to the natural environment and be least sensitive to the illumination changes that are most likely to occur (Aston et al., 2019). Alternatively, given the color distribution of lighting and shadows in natural scenes, the blue bias may emerge from the observer's tendency to infer blue tints as illuminants (Winkler et al., 2015). Although these explanations do not reconcile, they commonly imply an environmental account of the blue-yellow asymmetry: color vision may be adapted to natural spectra and expect blue illumination as a dominant feature in natural conditions.

In line with these interpretations, our observer model attributes the blue-yellow asymmetries at the behavioral level to a unimodal prior that peaks only at blue. Notably, the unimodal prior outperforms bimodal or uniform priors in predicting the perceptual biases (Fig. S7). Furthermore, our model is in line with the notion that perception, and in particular color perception, is shaped by the regularities of the sensory environment (Shepard, 1992), but also suggests an asymmetry in natural daylight. As Mollon (2006) has pointed out, the clear sky appears unique blue, which strongly indicates that the light of the sky, resulting from the fundamental physical process of Rayleigh scattering, might provide a stable reference to which color perception is anchored. Thus, in keeping with other Bayesian observer models (Ganguli and Simoncelli, 2010; Girshick et al., 2011), our results suggest that human perception internalizes the natural sensory statistics and incorporates prior knowledge into the processing of sensory information.

4 Methods

4.1 Participants

Six observers, (two males and four females), ranging in age from 23 to 56 years, participated in the experiments. All observers had normal or corrected-to-normal vision. Participants signed informed consent prior to the experiment and received a compensation of 10 Euros per hour. The study was conducted in accordance with the Declaration of Helsinki.

4.2 Stimuli

Stimuli were presented on a ViewPixx Lite 2000A display, calibrated by a PhotoResearch (Chatsworth, CA) PR-655 spectroradiometer, and controlled by a Radeon Pro WX 5100 graphics card. The screen resolution was set to 1920×1200 pixels at a refresh rate of 120 Hz.

Chromaticities of the stimuli were defined in an opponent cone-contrast color space, with the two coordinate axes corresponding to L-M and S cone contrast, respectively (Fig. 1a). In this color space, the distance from the center corresponds to chroma, the azimuth angle corresponds to hue, and the orthogonal third axis corresponds to luminance contrast. Cone contrasts were defined with respect to a neutral gray (106.7 cd/m^2, CIE [x, y] = [0.307, 0.314]). S-cone contrasts were scaled by a factor of 2.6, yielding approximately equally salient stimuli for all hues (Teufel and Wehrhahn, 2000). Individual perceptual isoluminance with respect to the reference gray was determined for 16 stimuli of different hues using heterochromatic flicker photometry (Kaiser and Boynton, 1996). From these data, an isoluminant plane in cone-opponent color space was calculated individually for each subject (Teufel and Wehrhahn, 2000).

Stimuli were generated using Psychopy 2020.1.2 (RRID:SCR_006571, Peirce et al., 2019) based on Python 3.6. Each stimulus presentation consisted of a reference stimulus, a comparison stimulus, and a gradient color bar as a reference hue sequence along the color space azimuth (Fig. 1a-b). Both reference and comparison stimuli consisted of arrays of 16 circular patches with diameters of 0.75° of visual angle, evenly spaced on a 4×4 grid extending $3° \times 3°$ of visual angle. Within each stimulus, the circular patches were randomly positioned on the grid of the array. The gradient color bar had an extent of $12° \times 1.5°$ of visual angle.

On a neutral gray background ($46.01° \times 29.68°$), a fixation cross ($0.6° \times 0.6°$) was presented at the screen center. The gradient color bar was presented on the upper part of the screen, 5° above the center. Reference and comparison stimuli were placed 3° left and right from the fixation cross, with the sides of reference and comparison stimulus assigned in a pseudo-random fashion such that their positions were balanced within each session.

All stimulus chromaticities were moderately saturated, with a cone contrast $c = 0.12$ with respect to the gray background. All stimuli and the gray background were isoluminant. For both reference

and comparison stimuli, the hues of the circular patches either were identical (low-noise stimulus) or were drawn from a uniform distribution with a range that was individually pre-determined for each subject (high-noise stimulus, see below).

The reference hues, θ_r, were regularly spaced along the azimuth of color space at 22.5°, 67.5°, 112.5°, 157.5°, 202.5°, 247.5°, 292.5°, and 337.5°. For each θ_r, the corresponding gradient color bar was filled by 45 evenly spaced hue angles from a uniform distribution with the boundaries [θ_r - 30°, θ_r + 30°].

4.3 Procedures

During the experiment, the participant sat in a dimly lit room and viewed the display binocularly from a viewing distance of 57 cm. A central fixation cross was displayed and participants had to maintain their eyes to the fixation throughout the entire trial.

Each trial started with the presentation of the gradient color bar to indicate the range of hues to be tested. After 500 ms, the reference stimulus and the comparison stimulus were presented simultaneously for 500 ms, followed by a 500-ms full-screen checkerboard pattern of random chromatic squares (1.3° × 1.2°) to prevent afterimages of the stimuli. Participants were instructed to compare the two stimuli and indicate whether they were in the same sequence as the hue sequence of the gradient color bar. Responses were given by pressing the up or down arrow key with the right hand. There was no time constraint on the response. Average response times were within one second.

For a given reference stimulus, its average hue and corresponding gradient color bar remained unchanged across trials, while the hue difference between the reference hue and the average hue of the comparison stimulus was adjusted by a 1-up/2-down staircase. Two same-noise conditions, low-noise versus low-noise (L-L) and high-noise versus high-noise (H-H), and one cross-noise condition, low-noise versus high-noise (L-H), were tested in the experiments. In the cross-noise condition, the reference stimulus was always the low-noise stimulus.

Each session consisted of 128 trials, in which the first 8 trials were warm-up trials with random hues and were excluded from further analysis. Within each session, two reference stimuli with 180° difference in their mean hue angles, were randomly interleaved. Prior to the formal experiment,

participants completed four practice sessions with feedback about the correctness of their responses given as text emojis (":)" or ":(" for correct and incorrect responses, respectively). Each participant performed 5760 trials divided into 24 conditions (8 reference stimuli x 3 comparison conditions) over 48 sessions.

As in the study by Girshick et al. (2011), prior to the main experiment, we determined each participant's level of hue variance for the high-noise stimulus such that H-H discrimination thresholds were roughly 2 times larger than L-L discrimination thresholds. A reference hue angle of 135° was used because thresholds had intermediate values for this hue. For each participant, we measured the discrimination threshold (d_L) of this stimulus in the L-L condition by fitting a cumulative Gaussian function to the psychophysical data. Next, we fixed the reference and comparison stimuli at 135° and 135°+$2d_L$, respectively, and used a 1-up/2-down staircase to adjust the amount of the hue noise of both stimuli. The noise level was chosen that yielded 75% correct performance according to the psychometric function. The determined noise levels (21.2°, 20.1°, 24.5°, 20.1°, 22.2°, and 24.4° for 6 subjects, respectively) were used for the high-noise stimuli in the main experiment.

4.4 Data Analysis

4.4.1 Estimation of perceptual variability and bias

The psychophysical data were analyzed separately for each subject. In addition, data for a hypothetical average observer were obtained by pooling all subjects' data. For the data of each condition, we fitted a cumulative Gaussian function using non-linear least square minimization with the Nelder-Mead algorithm (Gao and Han, 2012) and determined its mean, representing the point of subjective equality (PSE) and the standard deviation, representing the just noticeable difference (JND). PSE and JND thus reflected perceptual bias and perceptual variability, respectively.

4.4.2 Estimation of the measurement distributions and the likelihood functions

The measurement distribution is the conditional distribution $p(m|\theta)$ and corresponds to the likelihood of a sensory measurement m given a particular stimulus hue angle θ. For each stimulus θ

we estimated the measurement distribution as a von Mises distribution with a peak at θ and the concentration parameter κ_θ, thus

$$p(m|\theta) = \frac{e^{(\kappa_\theta \cos(\theta - m))}}{2\pi i_0(\kappa_\theta)}, \tag{1}$$

where $i_0(\kappa_\theta)$ is the modified Bessel function of order 0. The concentration parameter κ_θ represented the measurement noise and was converted from the corresponding perceptual variability $J(\theta)$ by $\kappa_\theta = J^{-2}(\theta)$. We estimated $J(\theta)$ by fitting each subject's same-noise JNDs as a sine mixture function of the hue angle θ:

$$J(\theta) = a(\sin\theta + \sin 2\theta) + b. \tag{2}$$

With periods of 180° and 360°, the two sine functions allow capturing both the periodicity and asymmetry of the JND patterns. The parameters a and b were estimated using non-linear least squares minimization.

To avoid any constraint of predefined shapes in estimating the likelihood, we adopted a sampling method using a pre-computed two-dimensional function (Girshick et al., 2011), where one dimension represented measurement distributions centered on particular stimulus hue angles θ, and the other dimension represented likelihood functions of θ given particular measurements m. Thus, drawing a single measurement from the measurement distribution results in a likelihood, which corresponds to a horizontal slice of the two-dimensional function.

4.4.3 Estimation of the priors

As the model of the prior we chose a von Mises function of hue angle θ,

$$p(\theta) = \frac{e^{(\kappa \cos(\theta - m))}}{2\pi i_0(\kappa)}, \tag{3}$$

which guaranteed that the prior had a period of 360° and an integral of 1. We estimated the prior by fitting the simulation of a Bayesian observer to the behavioral data under the cross-noise condition (Girshick et al., 2011). We assume that the Bayesian observer encodes with sensory noise and gives fluctuated measurements $m(\theta)$ during the repeated presentation of the same stimulus

θ. Each measurement leads to a likelihood function, which is multiplied by the prior to obtain a maximum a posterior (MAP) estimate at the decoding phase. The MAP estimate thus represents the observer's estimate $\hat{\theta}$ of the stimulus θ. Therefore, the measurement distribution of a stimulus results in a distribution of MAP estimates. The discrimination task was simulated by comparing two MAP estimate distributions according to signal detection theory (Green and Swets, 1966), yielding a single point on the simulated psychometric function.

To simulate psychometric functions for each participant, we simulated 1000 trials for each stimulus pair of the cross-noise comparison data. For every paired low-noise and high-noise stimuli, 1000 samples each were drawn from two corresponding measurement distributions with centers at m_L and m_H, and concentration parameters κ_L and κ_H, respectively. Each sample generated a likelihood function that combined with the prior and led to an estimate of the stimulus. The two distributions formed by the 1000 estimates each were compared, resulting in a response probability given the corresponding stimulus pair. We then obtained a model-generated psychometric function by fitting a cumulative Gaussian function to the simulated data.

We evaluated the prior model by computing the likelihood of the cross-noise data given the model-generated psychometric function. The optimal parameters of the prior for each subject were estimated by maximizing the overall likelihood. We performed bootstrapping on each subject's binary responses for each stimulus pair 100 times and estimated the priors given the bootstrapped data. The standard deviation of the 100 estimated priors was taken as the confidence interval of the optimized prior. We further assessed the model by comparing its performance with the performance of a model with a uniform prior. We evaluated the performance of the model by a normalized difference of log-likelihood to the model with the uniform prior,

$$L = \frac{L_{\text{est}} - L_{\text{uni}}}{L_{\text{raw}} - L_{\text{uni}}}, \tag{4}$$

where L_{est} and L_{uni} represent the log likelihoods of the models with the estimated prior and the uniform prior, respectively, and L_{raw} represents the log likelihoods of the raw psychometric fits. Thus, $L = 0$ corresponds to the model with a uniform prior and $L = 1$ corresponds to the raw psychometric fits. Given the difference in the degree of freedom between the estimated prior and

the uniform prior, we also calculated the Akaike Information Criterion (AIC) scores (Akaike, 1998) of the models.

To validate the unimodality of the prior model, we also modeled two priors as normalized sine functions with periods of 360° and 180°, respectively, and compared the prediction performance of the corresponding observer models consisting of these priors.

Acknowledgments

Supported by DFG (grant number RTG 2175 "Perception in Context and its Neural Basis") and Bernstein Center for Computational Neuroscience Munich. We thank all participants for participating in the experiments.

References

Akaike, H. Information theory and an extension of the maximum likelihood principle. In Parzen, E, Tanabe, K, and Kitagawa, G, editors, *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer New York, New York, NY, 1998.

Aston, S, Radonjic, A, Brainard, DH, and Hurlbert, AC. Illumination discrimination for chromatically biased illuminations: Implications for color constancy. *Journal of Vision*, 2019, 19(3):15, 2019.

Bosten, JM, Beer, RD, and MacLeod, DIA. What is white? *J. Vis.*, 2015, 15(16):5, 2015.

Boynton, RM, Nagy, AL, and Eskew, RT, Jr. Similarity of normalized discrimination ellipses in the constant-luminance chromaticity plane. *Perception*, 1986, 15(6):755–763, 1986.

Brainard, DH and Freeman, WT. Bayesian color constancy. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 1997, 14(7):1393–1411, 1997.

Brainard, DH, Longère, P, Delahunt, PB, Freeman, WT, Kraft, JM, and Xiao, B. Bayesian model of human color constancy. *Journal of Vision*, 2006, 6(11):1267–1281, 2006.

Danilova, MV and Mollon, JD. Parafoveal color discrimination: a chromaticity locus of enhanced discrimination. *Journal of Vision*, 2010, 10(1):4.1–9, 2010.

Delahunt, PB and Brainard, DH. Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision*, 2004, 4(2):57–81, 2004.

Emery, KJ, Volbrecht, VJ, Peterzell, DH, and Webster, MA. Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research*, 2017, 141:51–65, 2017.

Ernst, MO and Banks, MS. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 2002, 415(6870):429–433, 2002.

Ganguli, D and Simoncelli, EP. Implicit encoding of prior probabilities in optimal neural populations. *Advances in Neural Information Processing Systems*, 2010, 2010:658–666, 2010.

Gao, F and Han, L. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, 2012, 51(1):259–277, 2012.

Girshick, AR, Landy, MS, and Simoncelli, EP. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, 2011, 14(7):926–932, 2011.

Glasauer, S and Shi, Z. Individual beliefs about temporal continuity explain variation of perceptual biases. *Sci. Rep.*, 2022, 12(1):10746, 2022.

Green, DM and Swets, JA. *Signal detection theory and psychophysics.* Wiley New York, 1966.

Hansen, T, Walter, S, and Gegenfurtner, KR. Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision*, 2007, 7(4):2, 2007.

Juricevic, I and Webster, MA. Variations in normal color vision. V. Simulations of adaptation to natural color environments. *Visual Neuroscience*, 2009, 26(1):133–145, 2009.

Kaiser, PK and Boynton, RM. *Human Color Vision.* Washington: Optical Society of America, 1996.

Kellner, CJ and Wachtler, T. Stimulus size dependence of hue changes induced by chromatic surrounds. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2016, 33(3):A267–72, 2016.

Kersten, D, Mamassian, P, and Yuille, A. Object perception as Bayesian inference. *Annual Review of Psychology*, 2004, 55:271–304, 2004.

Klauke, S and Wachtler, T. "Tilt" in color space: Hue changes induced by chromatic surrounds. *Journal of Vision*, 2015, 15(13):17–17, 2015.

Knill, DC and Pouget, A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 2004, 27(12):712–719, 2004.

Körding, KP and Wolpert, DM. Bayesian integration in sensorimotor learning. *Nature*, 2004, 427 (6971):244–247, 2004.

Krauskopf, J and Gegenfurtner, K. Color discrimination and adaptation. *Vision Research*, 1992, 32 (11):2165–2175, 1992.

Macleod, DIA. Colour discrimination, colour constancy and natural scene statistics. In *Normal and Defective Colour Vision*. Oxford University Press, 2003.

Maule, J and Franklin, A. Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of Vision*, 2015, 15(4):6, 2015.

Maule, J, Witzel, C, and Franklin, A. Getting the gist of multiple hues: metric and categorical effects on ensemble perception of hue. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2014, 31(4):A93–102, 2014.

Mollon, J. Monge: The Verriest lecture, Lyon, July 2005. *Visual Neuroscience*, 2006, 23(3-4): 297–309, 2006.

Olkkonen, M, McCarthy, PF, and Allred, SR. The central tendency bias in color perception: effects of internal and external noise. *Journal of Vision*, 2014, 14(11), 2014.

Pearce, B, Crichton, S, Mackiewicz, M, Finlayson, GD, and Hurlbert, A. Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. *PLoS One*, 2014, 9(2):e87989, 2014.

Peirce, J, Gray, JR, Simpson, S, MacAskill, M, Höchenberger, R, Sogo, H, Kastman, E, and Lindeløv, JK. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 2019, 51(1): 195–203, 2019.

Radonjic, A, Pearce, B, Aston, S, Krieger, A, Dubin, H, Cottaris, NP, Brainard, DH, and Hurlbert, AC. Illumination discrimination in real and simulated scenes. *Journal of Vision*, 2016, 16(11):2, 2016.

Shepard, RN. The perceptual organization of colors: An adaptation to regularities of the terrestrial world? In Barkow, JH, editor, *The adapted mind: Evolutionary psychology and the generation of culture*, volume 666, pages 495–532. Oxford University Press, New York, 1992.

Shi, Z and Burr, D. Predictive coding of multisensory timing. *Curr Opin Behav Sci*, 2016, 8:200–206, 2016.

Shi, Z, Church, RM, and Meck, WH. Bayesian optimization of time perception. *Trends in Cognitive Sciences*, 2013, pages 1–9, 2013.

Stocker, AA and Simoncelli, EP. Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, 2006, 9(4):578–585, 2006.

Teufel, HJ and Wehrhahn, C. Evidence for the contribution of S cones to the detection of flicker brightness and red-green. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2000, 17(6):994–1006, 2000.

Tomassini, A, Morgan, MJ, and Solomon, JA. Orientation uncertainty reduces perceived obliquity. *Vision Research*, 2010, 50(5):541–547, 2010.

Virtanen, LS, Olkkonen, M, and Saarela, TP. Color ensembles: Sampling and averaging spatial hue distributions. *Journal of Vision*, 2020, 20(5):1, 2020.

von Helmholtz, H. *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln.* Voss, 1867.

Wachtler, T, Lee, TW, and Sejnowski, TJ. Chromatic structure of natural scenes. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2001, 18(1):65–77, 2001.

Wachtler, T, Sejnowski, TJ, and Albright, TD. Representation of color stimuli in awake macaque primary visual cortex. *Neuron*, 2003, 37(4):681–691, 2003.

Webster, J, Kay, P, and Webster, Ma. Perceiving the average hue of color arrays. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2014, 31(4):A283–92, 2014.

Webster, MA and Mollon, JD. Adaptation and the color statistics of natural images. *Vision Research*, 1997, 37(23):3283–3298, 1997.

Webster, MA, Miyahara, E, Malkoc, G, and Raker, VE. Variations in normal color vision. i. Cone-opponent axes. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2000a, 17(9):1535–1544, 2000a.

Webster, MA, Miyahara, E, Malkoc, G, and Raker, VE. Variations in normal color vision. II. Unique hues. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2000b, 17 (9):1545–1555, 2000b.

Webster, MA. The verriest lecture: Adventures in blue and yellow. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2020, 37(4):V1–V14, 2020.

Webster, MA, Webster, SM, Bharadwaj, S, Verma, R, Jaikumar, J, Madan, G, and Vaithilingham, E. Variations in normal color vision. III. Unique hues in indian and united states observers. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 2002, 19(10): 1951–1962, 2002.

Wei, XX and Stocker, AA. A Bayesian observer model constrained by efficient coding can explain'anti-Bayesian'percepts. *Nature Neuroscience*, 2015, 18(10):1509, 2015.

Wei, XX and Stocker, AA. Lawful relation between perceptual bias and discriminability. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114(38):10244–10249, 2017.

Weiss, D, Witzel, C, and Gegenfurtner, K. Determinants of colour constancy and the blue bias. *Iperception*, 2017, 8(6):2041669517739635, 2017.

Winkler, AD, Spillmann, L, Werner, JS, and Webster, MA. Asymmetries in blue-yellow color perception and in the color of 'the dress'. *Current Biology*, 2015, 25(13):R547–8, 2015.

Witzel, C and Gegenfurtner, KR. Categorical sensitivity to color differences. *Journal of Vision*, 2013, 13(7):1, 2013.