Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

Importance of hue: color discrimination of three-dimensional objects and two-dimensional discs

Laysa Hedjar,[1,*] Matteo Toscani,[2] and Karl R. Gegenfurtner[1]

[1] Abteilung Allgemeine Psychologie, Justus-Liebig-Universität-Gießen, Gießen, Germany
[2] Department of Psychology, Bournemouth University, Poole BH12 5BB, UK
*laysa.hedjar@psychol.uni-giessen.de

Received 16 October 2024; revised 29 January 2025; accepted 4 February 2025; posted 5 February 2025; published 3 March 2025

While flat, 2D stimuli have traditionally been used to measure color discrimination, our everyday interactions typically involve 3D objects. Here, we compare discrimination thresholds for rendered matte 3D objects and uniform discs. Participants performed a 4AFC odd-one-out task, where the odd stimulus reflectance differed in hue or chroma in four quadrants of DKL color space. Hue thresholds for 3D objects and 2D discs were equal, while object chroma thresholds were significantly higher, suggesting that hue is especially important for object discrimination. Chroma-to-hue threshold ratios were above 1 in all quadrants, particularly the bluish and orangish where a preponderance of natural object reflectances plot. This supports the idea that hue is also more important for the object colors we see most in our environment.

https://doi.org/10.1364/JOSAA.544380

1. INTRODUCTION

Estimates of how many colors humans can discriminate vary widely (e.g., [1–5]), but there is consensus that the discrimination ability is far from uniform across any modern color space. Early measurements by MacAdam [6] showed elliptical discrimination contours in the CIE 1931 xy chromaticity space, but numerous attempts to use these measurements to create a uniform color space have been unsuccessful (for an overview, see Ref. [7]). In the Munsell color space, Judd revealed through computations using Nickerson's formula [8] that for every just-noticeable chroma difference, one would find twice as many just-noticeable differences in hue. Terming this the "super-importance of hue," he concluded that this would make it impossible to create a uniform Euclidean color space [9]. Other studies showed that these theoretical considerations are indeed backed up by empirical data. Hue discrimination thresholds are smaller than chroma thresholds ([3,10–12], but see Ref. [13]) for the postreceptoral color spaces DKL [14] and MacLeod–Boynton [15].

Despite endeavors to create a perceptually uniform color space, there is little understanding of why color discrimination is non-uniform in the first place. Danilova and Mollon [10] hypothesized in their discussion that the "super-importance of hue" could be explained by correlated neural noise between two channels. Stimuli coded by two channels and giving rise to differences in chromatic contrast elicit greater noise correlation between the two channels than stimuli giving rise to differences in hue, leading to higher thresholds. Giesel *et al.* showed that, for chromatically varying textures or natural objects, thresholds mainly increased as a function of distance from the adaptation point, i.e., chroma, along the direction of chromatic variation [16,17] (see also Ref. [18]). Even when the chromatic variation was mainly along the hue direction, discrimination ellipses were still more elongated along the chroma direction [16,17].

Much of the experiments on color discrimination thresholds ask observers to differentiate between uniform discs or multi-colored patches. While these works have advanced our understanding of color, they are still far from the tasks that require color discrimination in our daily life. Choosing a specific fruit when grocery shopping, picking out matching socks from a load of laundry, or using the coloration of a patient's skin for a diagnosis all involve discriminating between objects which have distributions of colors with natural variation. Do the discrimination thresholds we find when using unnatural stimuli generalize to realistic objects, and can the differences in hue and chroma thresholds across color space be explained by the natural variation of objects?

The distribution of colors within natural objects varies systematically with respect to chroma and intensity, but is relatively stable in hue [19–25]. For an object with a uniform diffuse material across its surface, the variation in the light that reaches the eye from this object falls along the chroma and luminance dimensions of the color space due to shading. This arises from the interaction between the angle of incidence of the illumination beam and the surface normal of the object and follows Lambert's cosine law, such that the smaller the difference in angle between the two, the higher the chroma and luminance. While shading is important for 3D shape perception [26–30], for object discrimination hue is much more relevant. If color discrimination is driven by the colors of our environment, our reduced sensitivity to chroma changes may arise because chroma differences, largely due to shading, are not as relevant as hue differences. The visual system might have adapted to have increased sensitivity to hue differences in order to maximize object differentiation [11]. Moreover, we know that hue thresholds vary across regions of these color spaces: for the orangish [31] and bluish quadrants, hue thresholds are much smaller than chroma thresholds, but for purplish and greenish, the thresholds are more comparable [3,10–12,16,17]. This matches the well-known abundance of natural reflectances in the "orange" region of warm colors [32–35].

We examined whether hue and chroma discrimination thresholds of uniform discs are similar to those of rendered 3D objects. The light reflected from these objects varied naturally in chroma and luminance due to shading but little in hue. We found that the ratio of chroma-to-hue thresholds was elevated for the 3D objects. While hue thresholds for the rendered 3D objects and discs were approximately the same, chroma thresholds were significantly elevated for the objects, albeit by a small amount compared to the increase in chroma variation. Thus, the natural shading of the objects had no effect on hue discrimination, while shading elicits greater chroma variation within an object, which in turn elevates its chroma thresholds. Additionally, and in line with previous work, we also found that the ratio of chroma-to-hue thresholds was larger for colors in the orangish and bluish quadrants rather than purplish and greenish in DKL space; we saw this difference for both disc and 3D object stimuli.

2. METHODS

A. Definition of DKL Color Space

The axes of DKL color space were specified according to the formulas provided in Hansen and Gegenfurtner [36]. We defined the isoluminant plane in DKL in terms of the monitor at midgray, and all our stimulus colors were defined on this plane. We used a Konica Minolta CS2000A spectroradiometer to measure the spectrum of the R, G, and B channels of the monitor. We converted the spectra to xyY values using Judd-corrected color-matching functions. xyY coordinates of the RGB channels were: $R = (0.6840, 0.3136, 29.1504)$, $G = (0.2116, 0.7286, 80.9724)$, and $B = (0.1533, 0.0559, 7.1281)$. Thus, the DKL-to-RGB transformation matrix was as shown in Table 1. This was calculated using Smith and Pokorny [37] 2° cone fundamentals (with the Boynton [38] Z coefficient value). The nominal CIE 1931 (Judd-corrected) xyY values of the monitor at midgray were (0.3216, 0.3535, 58.63).

Table 1. DKL-to-RGB Transformation Matrix

	L + M	L − M	S − (L + M)
R	1	1	0.0676
G	1	−0.3619	−0.1124
B	1	0.0216	1

Since the scaling of the S-(L + M)-axis of DKL is not standardized [36], we scaled it such that detection thresholds at the adaptation point were approximately symmetrical. The scaling factor was determined using a detection task performed by five experienced observers including one of the authors. A uniformly colored disc could appear at one of four locations around the fixation dot at the center of the screen. The disc layout was the same as in the main experiment's control task (2.5° in diameter, 0.56° horizontal and vertical spacing between possible positions). The disc appeared for 500 ms and observers used a keyboard to indicate the location of the disc. Observers received auditory feedback as in the main experiment (click sound = correct; white noise = incorrect). For three of the five observers, the intertrial interval was 1500 ms. For the other two, it was 1000 ms. Thresholds were determined in eight hue directions around the adaptation point. The color of the disc could be 0–0.06 DKL units away from the adaptation point; we predefined 15 possible steps. For one observer (an author), 11 steps were defined between 0 and 0.04 units from the adaptation point. The color presented was determined by the adaptive staircase QUEST [39], with 70 trials per staircase. Trials for the eight staircases were interleaved and split into three blocks about 13 min each, which observers completed on the same day.

We fitted psychometric curves to the results using the psignifit toolbox [40]. Thresholds were defined at 62.5% correct (halfway between chance and 100% correct). Figure 1 shows the thresholds collected from the five observers along with an ellipse best fit to the average across observers. The ellipse is closely aligned to the axes, in agreement with previous results [3,12,16,17]. There is at best a hint of a slight tilt toward the negative diagonal, but this is considerably smaller than in the discrimination ellipses at the adaptation point reported by Bosten *et al.* [41]. One difference between these studies is that Bosten and colleagues presented their stimuli on a luminance pedestal of >38% contrast.

We calculated the scaling factor by averaging thresholds along the S-(L + M) and L-M axes across observers, and then dividing the mean S-(L + M) threshold by the mean L-M threshold. This resulted in a factor of 1.63. All further coordinates and distances in DKL color space are defined in this scaled space using multiples of detection thresholds as units.

All stimuli were presented on an Eizo ColorEdge CG2420 (10 bits/channel) monitor with a resolution of 1920 × 1200 (46.9° × 30.3°). The stimuli were gamma-corrected using a look-up table before being presented on the monitor. We used Psychtoolbox-3 on MATLAB R2019b as our presentation software. Observers sat 60 cm from the monitor.

B. Stimuli

The reference points for the discrimination task bisected the four quadrants of DKL space and lay 24.8 detection threshold units away from the origin (midgray of the monitor). Figure 2 shows the isoluminant plane along with the reference points for the discrimination thresholds and the directions in which the thresholds were measured. The colors from each quadrant are roughly: purplish (Q1), bluish

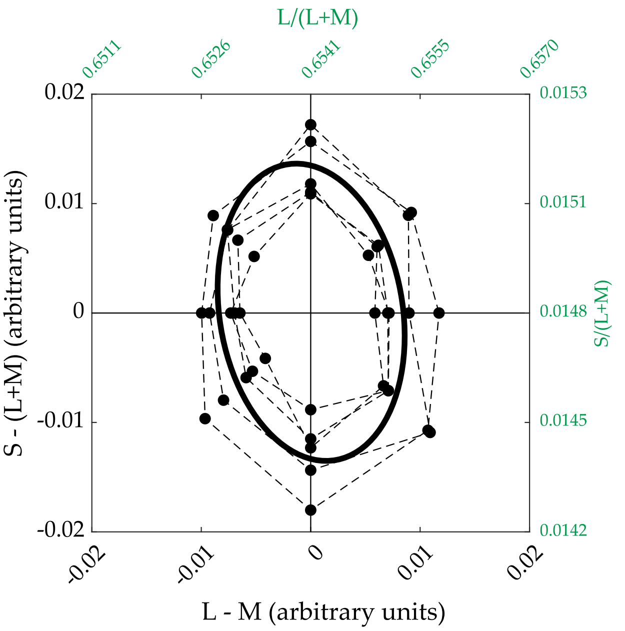

Fig. 1. Detection thresholds at the adaptation point (0,0). Black dots represent individual detection thresholds, and dashed lines connect thresholds for single observers. The thick black line indicates the best-fit ellipse for thresholds averaged across observers. The x and y axes are defined in arbitrary DKL units, where $[-0.5, 0.5]$ define the limits of the monitor gamut along this equiluminant plane. MacLeod–Boynton coordinates are also provided [15].

(Q2), greenish (Q3), and orangish (Q4). Their nominal y values are as follows: Q1 = (0.3275, 0.2907), Q2 = (0.2709, 0.3151), Q3 = (0.3123, 0.4510), and Q4 = (0.3862, 0.4024). They were equiluminant with the midgray of the monitor ($\sim$58.63 cd/m^2).

We used Mitsuba Renderer's (0.6 [42]) RGB-based rendering to create our stimuli. We rendered scenes of four 3D bumpy "blob" meshes floating in the air in a 2 × 2 configuration in a rectangular room with matte gray walls. The 3D blob objects, which we will frequently refer to as "blobs," had a diameter of $\sim$2.5° and were spaced $\sim$0.56° apart. A collimated light beam (RGB = 1; irradiance value = 3.15) placed overhead illuminated the blobs and the room in such a way that the blobs did not cast visible shadows on each other. Figure 2 depicts samples of the presented stimuli with reference colors from Quadrant 1. The 3D blobs had a matte material defined using the Ward BRDF model [43,44] implemented in Mitsuba. The Ward (or Ward-Dür) model is used to model anisotropic reflection. Its diffuse component specifies the reflectance component of the object which follows Lambert's law—that is, which reflects light in all directions regardless of the angle of incidence of the illuminant light. The specular component describes the reflectance component of the object which reflects light depending on the angle of incidence and the viewing angle [22,24,25]. Since we do not want our objects to have any specularity, we set the specular reflectance component to 0. The RGBs of the diffuse components of the blobs were defined in DKL space and then transformed to RGB. Three of the four blobs had the same diffuse RGB value (reference) while one (odd) was shifted from 0 to 12.4 detection threshold units either in hue or in chroma from the reference RGB (rendered in 15 possible steps). Chroma steps were shifted in the positive and negative radial directions, while hue steps were shifted along the tangents of the hue circle at the reference point (see Fig. 2).

The odd blob could be in one of four positions. We rendered four instances of each color step in each of these four positions, for a total of 16 possible scenes per step. For every trial, we randomly chose the position of the odd blob and randomly sampled from one of these four instances. No two blobs within a trial had the same 3D shape outline and shading profile. We ensured this by sampling from 64 spatial rotations of two blob meshes and checking that no two blobs within a scene were ever alike. We did this because we did not want observers to make their judgment by simply comparing the same point on each blob.

Because each 3D blob has a different shading profile, the distribution of pixels elicited by two blobs with the same diffuse-component RGB value was not necessarily the same. Each distribution was dependent on the interaction between the

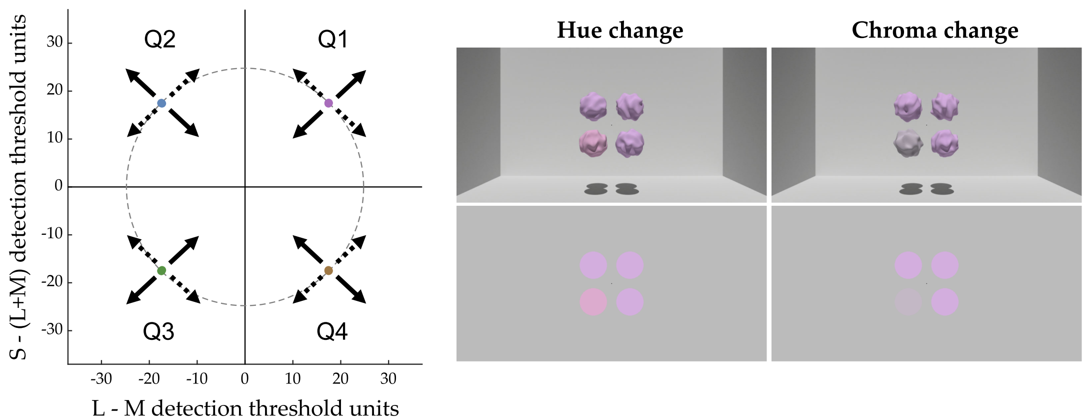

Fig. 2. Reference chromaticities and directions in which thresholds were measured, along with example stimuli. (Left) The reference colors and directions in which thresholds were measured are plotted on the isoluminant plane in DKL. The colored dots represent the reference points used during the experiment. Thresholds were measured in the positive and negative chroma directions (solid arrows) and the tangents of the clockwise and counterclockwise hue directions (dashed arrows). The hue circle is plotted as a light-gray dashed line. (Right) Example stimuli from Quadrant 1. The left images show a stimulus with the bottom left item shifted clockwise in hue, the right images in decreasing chroma. The top row depicts 3D blob stimuli and the bottom row 2D control disc stimuli. Differences are enhanced for visibility.

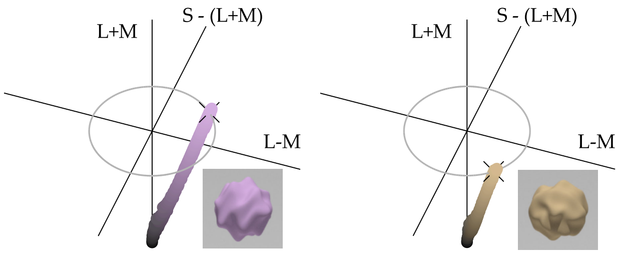

Fig. 3. Distribution of pixels reflected from sample 3D blobs defined by the RGBs of the reference colors from Quadrants 1 (left) and 4 (right). Blobs are rendered using the Mitsuba renderer (see text for details). The gray ring defines points of equal chroma along the isoluminant plane. A black X marks the DKL value given to the renderer to define each blob's diffuse component.

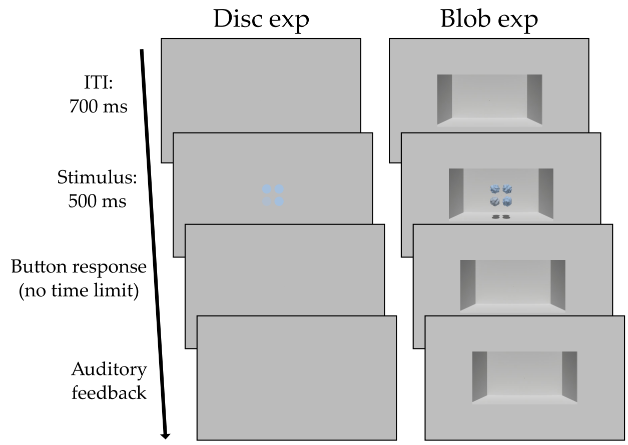

Fig. 4. Procedural layout of experiment. The stimulus was presented for 500 ms. Once the stimulus disappeared, observers used the keyboard to indicate in which location they believed the odd-one-out was. After their response, they received auditory feedback on their performance. There was a 700 ms intertrial interval between trials, during which we presented an empty room (3D blob stimuli) or a uniform gray screen (control 2D disc stimuli).

blob's shape (its concavity and convexity) and the angle of incidence of the light. Differences in these largely affected the chroma and luminance distributions of the pixels. Figure 3 plots the distribution of pixels emitted by a sample blob of the reference colors of Quadrants 1 and 4. In accordance with Lambert's cosine law, the distribution varies in chroma and luminance. Pixels with the highest chroma and luminance (on average, the top 4% brightest pixels) roughly corresponded to the DKL values plotted in Fig. 2. We chose to retain the natural variation in these blobs in order to keep their renderings as realistic as we could. However, this meant that different blobs defined by the same RGB value could have slightly different pixel distributions, and, therefore, different maximum chroma and luminance values. We attempted to quantify this by looking at the mean chroma of the top 4% brightest pixels of each rendered blob and calculating the range of means across blobs with the same diffuse component RGB. The range of such chroma means was around 4%–6% of the nominal chroma of the blobs. For the reference blobs at a chroma of 24.8 detection threshold units, this amounts to 0.99–1.49 threshold units. We consider this variation in our assessment of hue and chroma thresholds.

Because interreflections around the room and among the objects will lead to differences in the reflected light from each 3D blob depending on its position in the 2×2 configuration, we rendered each blob alone either in the top left or top right position and then cut and pasted it into its appropriate position in a prerendered empty gray room. To get the shadows in the scene, we rendered a room with four white spheres and cut and paste their shadows into the prerendered empty gray room. The walls of the room were also matte as defined by the Ward BRDF model, with the RGB of the diffuse component set to 0.603 (RGBs could range from 0 to 1). The reflected light that the renderer calculates depends not only on the RGBs of the object reflectances, but also on the intensity of the light source and the physical layout of the objects. Thus, the gray value of the room and the intensity of the light source were carefully chosen so that (1) the average of the immediate surround of the blob configuration (approximately a $10.3° \times 10.3°$ region) was the midgray of the monitor, and (2) the top 4% brightest points of the blobs were close to the RGB values of their diffuse components.

We note that the back wall of the room gets brighter as one moves toward the floor; this means that the immediate surround of the bottom two blobs was slightly brighter than that of the top two (roughly: bottom = 62.1 cd/m^2, top = 55.1 cd/m^2). We performed analyses to see if participants' calculated thresholds were different when they believed the blob was in the top row versus the bottom row and found no significant differences (see Appendix A).

The size of the final rendered scene was $25.4° \times 12.9°$ and always appeared in the center of the screen. We set the rest of the screen to RGB = 0.518 (60.73 cd/m^2), such that the average pixel value across the screen when the empty room was visible was the midgray of the monitor. All colors were gamma-corrected before presentation using a look-up table.

Control stimuli consisted of four uniform discs approximately the same size and position as the 3D blob stimuli (diameter = 2.5°, spacing = 0.56°) and presented on a uniform background set to the midgray of the monitor. Their colors were set to the RGB values used for rendering, which roughly correspond to the colors of the brightest pixels of the blobs.

C. Procedure

Each stimulus was displayed for 500 ms. After the stimulus disappeared from the screen, observers indicated using the keyboard which of the four stimuli was the "odd-one-out" in terms of color only. They were told to ignore differences in shape. Observers were given auditory feedback on their performance (click sound = correct; white noise = incorrect). During the intertrial interval (700 ms), the empty background scene (i.e., gray walls only) was presented during the blob blocks, and a uniform gray background was presented during the control blocks. Figure 4 outlines the procedural layout of the experiment.

The stimuli were presented according to the adaptive staircase method QUEST [39]. Observers completed 70 trials for each condition's staircase for a total of 2240 trials: four quadrants × four directions (two hue, two chroma) × two stimulus types (blobs and discs). Trials were blocked by stimulus type; all other conditions were interleaved. The trials were split into sessions. Observers with odd participant ID numbers completed sessions in the following order: disc, blob, disc, blob. Observers with even ID numbers completed sessions in the reverse order. Each session was split into three blocks of approximately 10 min each, between which observers could take self-paced breaks. One observer completed all sessions in one day; all others completed them over two days for a total of 2.5 h.

Observers completed three practice trials at the beginning of each session. They could repeat the practice trials if they were still unsure how to perform the task, but none of the observers required this. Each block began with 30 discarded trials. These trials allowed observers to adapt to the midgray of the monitor and also to grow accustomed to the difficulty of the task, but were not used in the analysis. The stimuli for the first 12 of these trials were presented at 1500 ms, the next six at 1000 ms, and the last 12 at 500 ms. These discarded trials merged seamlessly with the experimental trials.

We wanted to ensure that the center of the screen was at eye level for each observer, such that observers were not viewing the monitor at an angle. When the observers were first seated in the experimental room, we hung a taut string in front of the monitor which was level with a central fixation cross on the screen. We asked each observer to adjust the chinrest such that when they comfortably looked at the screen, the taut string fell in line with the fixation cross. During the experiment, a fixation dot was always present on the screen. We asked observers to keep their eyes on the fixation dot during stimulus presentation but told them they could blink during the intertrial intervals.

D. Participants

Twelve naïve observers (10 female) completed the task. Observers' ages ranged from 23 to 35 years old with a mean age of 27.8 years. All observers gave informed consent and had normal color vision as assessed by the Ishihara Color Vision Test [45].

E. Analyses

We fitted psychometric functions to the individual data using the *psignifit* toolbox for MATLAB [40]. The lower limit of the function was 25% correct (chance level for 4AFC) and the upper limit was 100% correct, so we defined discrimination thresholds at 62.5% correct. All statistical analyses were performed in the R programming environment [46]. We fitted linear mixed-effects models (LMMs) using the *nlme* package [47]. Best model fits were determined using the maximum-likelihood method; we used the structure of the best model and fitted the data using the restricted maximum-likelihood method for all analyses. ANOVAs were run using the built-in R package *stats*. All *post hoc* contrasts were performed using the *emmeans* package [48] with Bonferroni-corrected p-values.

3. RESULTS

Figure 5 plots the JNDs of all individuals for the 3D blob and the control disc stimuli on the equiluminant plane of DKL. We can see that Q2 (bluish) and Q4 (orangish) thresholds are more elliptical than Q1 (purplish) and Q3 (greenish) for both discs and blobs, supporting previous research [3,10–12,16,17].

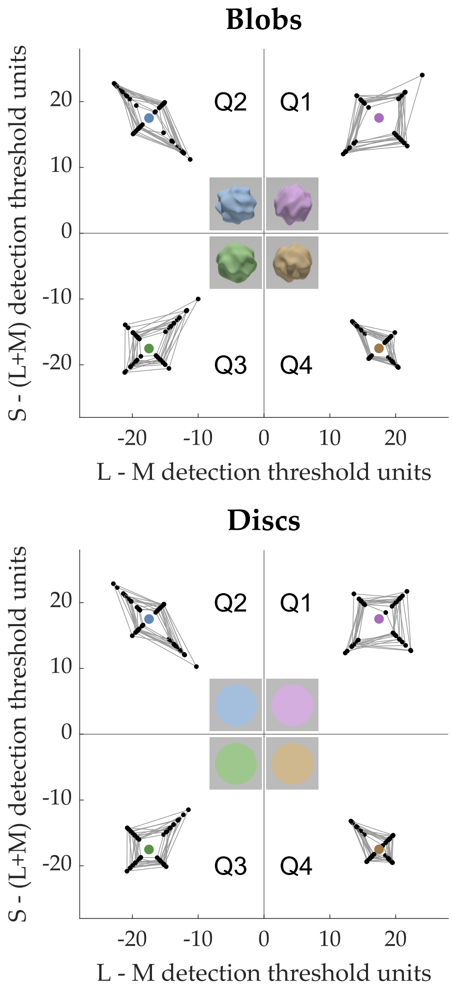

Fig. 5. JNDs of all participants for each condition, plotted in DKL color space, with the four reference points as dots color-coded by quadrant color. Each black dot is an individual observer's JND; JNDs for a given observer are connected with thin gray lines. Results for the 3D blobs are plotted in the top, for 2D control discs in the bottom.

Figure 6 plots the JNDs side-by-side. One can see that Q4 thresholds are the smallest among the four quadrants, and Q1 the largest. Hue thresholds between blob and disc stimuli are nearly identical, while chroma thresholds increase significantly for the blob stimuli.

We fitted a linear mixed-effects model with *stimulus type*, *color dimension*, and *quadrant* as fixed-effects factors and *observer* as a random-effects factor. We compared several models and found a best fit for a model with interactions between *stimulus type* and *color dimension* and between *quadrant* and *color dimension* included. To meet model assumptions, we used the log of the JND as the dependent variable. We performed a three-way ANOVA on the model. A main effect was present for all factors: 3D blob thresholds were higher than disc thresholds [$F(1, 363) = 9.25$, $p = 0.003$], chroma thresholds were larger than hue thresholds [$F(1, 363) = 333.1$, $p < 0.001$], and there was a significant difference between quadrants [$F(3, 363) = 72.1$, $p < 0.001$]. Pairwise contrasts (Bonferroni-corrected) revealed that thresholds in Q1 (purplish) were larger than in all other quadrants ($p < 0.001$) and

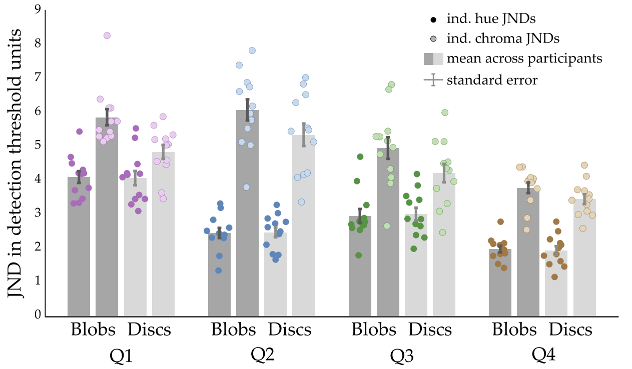

Fig. 6. JNDs in detection threshold units plotted on the y axis for each condition. Each data point represents an individual observer's JND, jittered slightly along the x axis for visibility. Data are color-coded by quadrant. Dots in line with the dark-gray bars represent blob stimuli, dots in line with the light-gray bars represent discs. Darker dots represent hue JNDs and lighter dots chroma JNDs. The height of the bars represents the mean JND across participants, with error bars ±1 SEM.

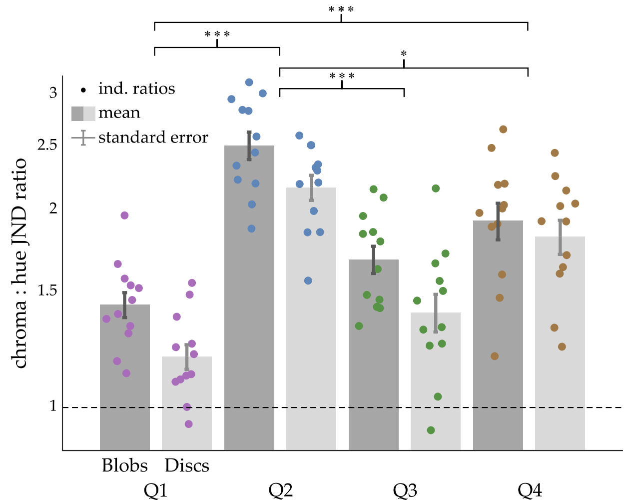

Fig. 7. Chroma-to-hue JND ratios per quadrant and stimulus type. Ratios are plotted on a logarithmic scale. Each data point is an individual observer's ratio. Ratios are color-coded by quadrant. Data points are jittered along the x axis for visibility. Gray bars (dark gray = blobs, light gray = discs) represent the geometric mean of the ratios with ±1 standard error [50]. The dashed line at $y = 1$ marks the point at which hue and chroma JNDs would be equal. Note that the y axis is log-scaled. Results from the pairwise comparisons between quadrants are indicated with asterisks between quadrants: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

Q4 (orangish) thresholds were smaller than all other quadrants ($p < 0.001$).

We also found a significant interaction between *stimulus type* and *color dimension* [$F(1, 363) = 8.14$, $p = 0.005$]. Blob and disc hue thresholds were not significantly different ($p = 1$), but chroma thresholds for the blobs were significantly larger than for discs ($p < 0.001$). Average raw chroma JNDs were 4.46 detection threshold units for discs and 5.20 for blobs. Another significant interaction was found between *quadrant* and *color dimension* [$F(3, 363) = 19.3$, $p < 0.001$]. Hue thresholds for Q1 (purplish) were larger than in all other quadrants ($p < 0.001$), while Q1's chroma thresholds were larger than all but Q2's (bluish; $p = 1$). Hue thresholds for Q2 were significantly larger than Q4's (orangish; $p = 0.001$) but smaller than Q3's (greenish; $p = 0.019$). Q2's chroma thresholds, however, were significantly larger than both Q3's and Q4's ($p < 0.001$).

We particularly wanted to compare ratios between chroma and hue for each condition. Figure 7 plots the chroma-to-hue ratios of each stimulus type and quadrant for each observer. We applied a two-way ANOVA with *stimulus type* and *quadrant* as main effects but with no interaction [49] to our data. Chroma-to-hue JNDs ratios for 3D blobs were significantly larger than for discs [$F(1, 187) = 7.76$, $p = 0.006$], and there was a significant difference across quadrants [$F(3, 187) = 18.4$, $p < 0.001$]. Pairwise contrasts indicated that ratios for Q2 (bluish) were larger than for all other quadrants: Q1 [purplish; $t(187) = 6.88$, $p < 0.001$], Q3 [greenish; $t(187) = 5.40$, $p < 0.001$], and Q4 [orangish; $t(187) = 2.78$, $p = 0.04$]. Ratios for Q1 were also significantly smaller than for Q4 [$t(187) = 4.10$, $p < 0.001$], but not for Q3 [$t(187) = 1.49$, $p = 0.84$]. Ratios for Q3 and Q4 were not significantly different [$t(187) = 2.62$, $p = 0.057$].

4. DISCUSSION

Here, we explored whether color discrimination thresholds for uniformly colored 2D discs are similar to those for rendered 3D blobs which elicit distributions of colors. These distributions varied in chroma and intensity, but very little in hue. We found that hue thresholds between blobs and discs were nearly identical. This suggests that the variation in luminance and chroma due to shading across an object has no effect on our ability to discriminate its hue. Observers were able to identify the odd 3D blob when it differed in hue just as well as with uniformly colored 2D discs, regardless of the different blob shapes, shading profiles, and relationships between luminance, chroma, and hue. This highlights the importance of hue in object identification and segmentation [20,21,51].

While chroma thresholds for the 3D blobs were elevated compared to the 2D discs, this increase was relatively modest (on raw thresholds, Cohen's $d = 0.40$) given that there was very large chroma overlap between the distributions of chroma-shifted blobs (Fig. 8, left): two blobs whose colors are just-noticeably differentiable along the chroma dimension have a chroma overlap of at least 74%. Because of this elevation in chroma thresholds, we also found that chroma-to-hue ratios for blobs were significantly larger than for discs. We note, though, that the distributions are much more differentiable if one also considers their luminance and how it relates to their chroma: in 3D DKL space, which includes the luminance dimension, the distributions do not overlap (Fig. 8, right). The importance of the interaction between chroma and luminance is supported by work from Hedjar *et al.* [52]. They created pairs of stimuli with equal chroma and luminance distributions but opposing relationships—one positive and one negative—and asked observers to judge which was more saturated. Observers consistently chose the one with the positive luminance-chroma relationship as more saturated, despite the other having larger nominal values for mean and max saturation (see note [53]

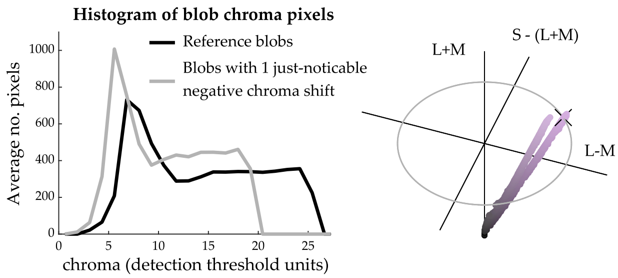

Fig. 8. Distribution of blob pixels. (Left) Histogram of chroma pixels averaged across all reference blobs (black line) and of blobs approximately 1 Q1 chroma JND away from the reference (gray line). (Right) 3D plot of the distribution of pixels in DKL reflected from two sample Q1 blobs differing by 1 chroma JND. One can see that the distributions overlap heavily along the chroma dimension, but that these distributions can be separated when luminance is taken into account.

and [54–56]). With respect to object judgments, chroma and luminance are interpreted jointly.

We also know that observers use the brightest regions of an object to make lightness judgments [20,57–60]. It is unclear whether we use the same regions to make judgments about object color in general; however, Hedjar *et al.* [52] show that saturation judgments are better predicted by the nominal saturation values of the brighter parts of the object. It may be that observers largely focus on the brightest regions of objects when discriminating between them. Given that chroma and luminance are positively correlated for blob pixel distributions, the large chroma overlap for chroma-shifted blobs occurs mostly in the darker regions of the blobs. It would therefore be reasonable for observers to ignore the darker regions and base their judgments on the brighter ones, which make the blobs more easily differentiable. However, if this is the case we still do not know how much observers might weigh brighter regions over darker.

We note additionally that variations in the 3D shape of our blobs led to slightly different chroma distributions among blobs defined by the same nominal color (up to 6% for the brightest 4% of pixels). We can expect that the higher blob chroma threshold can partially be attributed to this variation. However, we do not know exactly how observers weigh the distribution of pixels in each blob to make their color discrimination judgment—they may rely most on the top 50% of pixels or only on the top 1%. The higher 3D blob chroma threshold is most likely due to a combination of this blob distribution noise and the fact that the blobs elicit a distribution of colors compared to discs. From our experiment, we cannot separate these two possible sources of higher blob chroma thresholds, but our results clearly show that the effect of increased chromatic distribution has a disproportionally smaller effect on thresholds, similar to previous work on textures and discs [16,17].

Supporting previous literature, our results show the differences between hue and chroma thresholds across quadrants: purplish and greenish colors have the lowest chroma-to-hue ratios while bluish and orangish have the highest (although ratios for greenish and orangish quadrants were not significantly different). In particular, raw thresholds were highest for purplish colors and lowest for orangish colors, although bluish colors had the highest chroma-to-hue ratios. We also found that the threshold difference between blobs and discs was similar across quadrants, indicating that the natural distribution of light reflected from an object does not affect our discrimination of objects differently depending on the region of the color space. The results from this experiment support the theory that human color discrimination is shaped by object discrimination in the environment—the noise added by the chromatic distribution of objects makes little difference in our discrimination abilities.

Our aim was to examine the extent to which hue and chroma color discrimination of uniform discs (a task unlikely for the human visual system to encounter in the real world) can generalize to hue and chroma color discrimination of realistic objects (more likely for our visual system to do in daily life). We chose to limit the differences between the experimental and the control experiment to a handful of factors: we used a set of rendered 3D objects instead of uniform 2D discs, and we presented them inside a simple gray room instead of on a uniform gray background. Thus, we intentionally did not apply several optical principles that would normally occur in the real world. For example, we eliminated interreflections both between objects and between objects and the background, including the effect of object position in relation to the background (while retaining the natural interreflections *within* objects and background). We also positioned the four objects in a way that one is unlikely to encounter in the real world—floating in the air above each other. These changes limit the ecological validity of our stimulus; we know that interreflections play an important role in interpreting a scene [61,62] and that there can be drastic differences between the light field caused by the light source alone and the light field caused by the light source plus interreflections [63].

5. CONCLUSION

In conclusion, our findings suggest that the importance of hue for the discrimination of rendered 3D objects is even higher than for uniform discs. This stems from an increase in chroma thresholds for the 3D objects, likely due to the objects having greater chroma variation, while hue thresholds were unaffected. However, chroma thresholds between objects and discs were surprisingly similar in relation to their massive difference in chroma variation. Both types of stimuli elicited higher chroma than hue thresholds despite the absence of chroma variation in the discs. This suggests that the importance of hue in discrimination is mainly driven by object color statistics. We observed a higher chroma-to-hue threshold ratio for colors in the orangish and bluish quadrants of DKL space, as opposed to the purplish and greenish regions, for both 2D discs and 3D objects. This agrees with the theory that color discrimination is tuned to the natural color statistics of our environment.

APPENDIX A

A.1. Examination of Position Bias for 3D Blob Stimuli

In our experiment, we tested discrimination thresholds for rendered 3D "blob" stimuli which were presented on the screen in a 2×2 configuration in a rendered room. As noted in Section 2, the background gray behind the top row of blob stimuli was slightly darker than the background behind the bottom row due to interreflections between the walls of the room, leading to a small difference in luminance-contrast between stimulus and background for the two rows. We wanted to examine whether this resulted in a bias in the thresholds for the blob stimuli. We split the data into two sets based on whether observers believed the odd blob was in the top row versus the bottom row. We calculated thresholds for both datasets as described in Section 3 and compared thresholds. We compared several linear mixed-effects models and fitted one with *quadrant*, *color dimension*, *direction* (positive or negative chroma shift, increasing or decreasing L-M hue shift [64]), and *row response* as fixed-effects factors, including all interactions in the model, and with *observer* as a random-effects factor with random slopes for *color dimension* and *direction*. Performing a three-way ANOVA on this model, we found no significant effect of *row response* [$F(1, 363) = 0$, $p = 1.0$]. However, we did find that the interaction between *row response* and *direction* was significant [$F(1, 341) = 8.58$, $p = 0.004$] as well as the three-way interaction between *row response*, *direction*, and *color dimension* [$F(1, 341) = 13.39$, $p < 0.001$] and a four-way interaction across all factors [$F(3, 341) = 2.68$, $p = 0.047$]. Across these interactions, only one meaningful pairwise contrasts involving *row response* was significant: thresholds for bottom-row-chosen trials with decreasing chroma were larger than top-row-chosen [$t(341) = 4.02$, $p = 0.002$, estimated marginal mean difference $= 0.82$ detection threshold units]. While we did compute thresholds for each shift direction, we averaged thresholds between these directions in our main analyses. Given this, and that all significant model terms with *row response* interacted with *direction*, we believe that the effect of the background luminance difference between the top and bottom row blobs did not have a meaningful impact on our results. Nonetheless, we do not dismiss the role of background luminance on discrimination thresholds for our 3D blobs, especially since chroma and luminance were correlated. An experiment explicitly designed to investigate this would better help us understand its effect.

Funding. European Research Council (Advanced Grant Color 3.0 (884116)).

Acknowledgment. The authors would like to thank Maike Wilhelms for help on pilot versions of this study.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in Ref. [65].

REFERENCES AND NOTES

1. J. Koenderink, A. Van Doorn, and K. Gegenfurtner, "Graininess of RGB-display space," i-Perception **9**, 204166951880397 (2018).
2. J. Koenderink, A. Van Doorn, and K. Gegenfurtner, "Colour order," i-Perception **10**, 2041669519872516 (2019).
3. J. Krauskopf and K. Gegenfurtner, "Color discrimination and adaptation," Vis. Res. **32**, 2165–2175 (1992).
4. J. M. M. Linhares, P. D. Pinto, and S. M. C. Nascimento, "The number of discernible colors in natural scenes," J. Opt. Soc. Am. A **25**, 2918–2924 (2008).
5. M. R. Pointer and G. G. Attridge, "The number of discernible colours," Color Res. Appl. **23**, 52–54 (1998).
6. D. L. MacAdam, "Visual sensitivities to color differences in daylight*," J. Opt. Soc. Am. **32**, 247–274 (1942).
7. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. (Wiley, 1982), Vol. **40**.
8. D. Nickerson, "The specification of color tolerances," Text. Res. **6**, 505–514 (1936).
9. D. B. Judd, "Ideal color space redefined," Palette **31**, 23–29 (1969).
10. M. V. Danilova and J. D. Mollon, "Superior discrimination for hue than for saturation and an explanation in terms of correlated neural noise," Proc. Biol. Sci. **283**, 20160164 (2016).
11. M. V. Danilova and J. D. Mollon, "Symmetries and asymmetries in chromatic discrimination," J. Opt. Soc. Am. A **31**, A247–A253 (2014).
12. L. Hedjar, M. Toscani, and K. R. Gegenfurtner, "Importance of hue: the effect of saturation on hue-chroma asymmetries," J. Opt. Soc. Am. A **42**, B305–B312 (2025).
13. S. E. Regan, R. J. Lee, D. I. A. MacLeod, *et al.*, "Are hue and saturation carried in different neural channels?" J. Opt. Soc. Am. A **35**, B299–B308 (2018).
14. A. M. Derrington, J. Krauskopf, and P. Lennie, "Chromatic mechanisms in lateral geniculate nucleus of macaque," J. Physiol. **357**, 241–265 (1984).
15. D. I. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance," J. Opt. Soc. Am. **69**, 1183–1186 (1979).
16. M. Giesel, T. Hansen, and K. R. Gegenfurtner, "The discrimination of chromatic textures," J. Vis. **9**(9), 11 (2009).
17. T. Hansen, M. Giesel, and K. R. Gegenfurtner, "Chromatic discrimination of natural objects," J. Vis. **8**(1):2.1–19 (2008).
18. S. F. te Pas and J. J. Koenderink, "Visual discrimination of spectral distributions," Perception **33**, 1483–1497 (2004).
19. R. Ennis, F. Schiller, M. Toscani, *et al.*, "Hyperspectral database of fruits and vegetables," J. Opt. Soc. Am. A **35**, B256–B266 (2018).
20. M. Giesel and K. R. Gegenfurtner, "Color appearance of real objects varying in material, hue, and shape," J. Vis. **10**(9), 10 (2010).
21. C. Witzel and H. Dewis, "Why bananas look yellow: the dominant hue of object colours," Vis. Res. **200**, 108078 (2022).
22. S. A. Shafer, "Using color to separate reflection components," Color Res. Appl. **10**, 210–218 (1985).
23. J. Mollon, "Monge: The Verriest Lecture, Lyon, July 2005," Visual Neurosci. **23**, 297–309 (2006).
24. G. Monge, "Mémoire sur quelques phénomènes de la vision," Ann. Chim. **3**, 131–147 (1789).
25. H.-C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," J. Opt. Soc. Am. A **3**, 1694–1699 (1986).
26. J. J. Koenderink, *Solid Shape* (MIT, 1990).
27. B. Kunsberg, D. Holtmann-Rice, E. Alexander, *et al.*, "Colour, contours, shading and shape: Flow interactions reveal anchor neighbourhoods," Interface Focus **8**, 20180019 (2018).
28. M. S. Langer and H. H. Bülthoff, "A prior for global convexity in local shape-from-shading," Perception **30**, 403–410 (2001).
29. V. S. Ramachandran, "Perception of shape from shading," Nature **331**, 163–166 (1988).
30. A. J. van Doorn, J. J. Koenderink, and J. Wagemans, "Light fields and shape from shading," J. Vis. **11**(3), 21 (2011).
31. We use "orangish," "bluish," "purplish," and "greenish" because the terms roughly represent the appearance of stimuli in these quadrants of DKL color space. We do not claim that observers actually use these terms to describe the colors within those quadrants.
32. J. Koenderink, A. van Doorn, and K. Gegenfurtner, "RGB colors and ecological optics," Front. Comput. Sci. **3**, 630370 (2021).
33. I. A. Rosenthal, S. Ratnasingam, T. Haile, *et al.*, "Color statistics of objects, and color tuning of object cortex in macaque monkey," J. Vis. **18**(11), 1 (2018).
34. T. Morimoto, C. Zhang, K. Fukuda, *et al.*, "Spectral measurement of daylights and surface properties of natural objects in Japan," Opt. Express **30**, 3183–3204 (2022).
35. T. Hansen and B. R. Conway, "The color of fruits in photographs and still life paintings," J. Vis. **24**(5), 1 (2024).
36. T. Hansen and K. R. Gegenfurtner, "Higher order color mechanisms: evidence from noise-masking experiments in cone contrast space," J. Vis. **13**(1), 26 (2013).
37. V. C. Smith and J. Pokorny, "Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm," Vis. Res. **15**, 161–171 (1975).

38. R. M. Boynton, *Human Color Vision* (Holt, Rinehart & Winston, 1979).
39. A. B. Watson and D. G. Pelli, "Quest: a Bayesian adaptive psychometric method," Percept. Psychophys. **33**, 113–120 (1983).
40. H. H. Schütt, S. Harmeling, J. H. Macke, *et al.*, "Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data," Vis. Res. **122**, 105–123 (2016).
41. J. M. Bosten, R. D. Beer, and D. I. A. MacLeod, "What is white?" J. Vis. **15**(16), 5 (2015).
42. W. Jakob, "Mitsuba 0.6," (2010).
43. D. Geisler-Moroder and A. Dür, "A new ward BRDF model with bounded Albedo," Comput. Graph. Forum **29**, 1391–1398 (2010).
44. G. J. Ward, "Measuring and modeling anisotropic reflection," Comput. Graph. **26**, 265–272 (1992).
45. S. Ishihara, *Ishihara's Tests for Colour Deficiency*, 38th ed. (Kanehara and Co., Ltd., 2018).
46. R. C. Team, "R: A Language and Environment for Statistical Computing," R Foundation for Statistical Computing (2023).
47. J. Pinheiro, D. Bates, and R. C. Team, "Nlme: Linear and nonlinear mixed effects models," (2023).
48. R. V. Lenth, "Emmeans: Estimated marginal means, aka least-squares means" (2023).
49. Including an interaction term in the model neither improved the model fit nor changed the results, and fitting a linear mixed-effects model with *observer* as a random-effects factor had no effect either.
50. N. Norris, "The standard errors of the geometric and harmonic means and their application to index numbers," Ann. Math. Stat. **11**, 445–448 (1940).
51. T. Hansen and K. R. Gegenfurtner, "Independence of color and luminance edges in natural scenes," Visual Neurosci. **26**, 35–49 (2009).
52. L. Hedjar, M. Toscani, and K. R. Gegenfurtner, "Perception of saturation in natural objects," J. Opt. Soc. Am. A **40**, A190–A198 (2023).
53. Saturation is defined as chroma divided by lightness in CIELUV color space [53], but can be also applied to other color spaces such as CIELAB and DKL [54,55].
54. M. D. Fairchild, *Color Appearance Models*, 3rd ed. (Wiley, 2013).
55. F. Schiller, M. Valsecchi, and K. R. Gegenfurtner, "An evaluation of different measures of color saturation," Vis. Res. **151**, 117–134 (2018).
56. F. Schiller and K. R. Gegenfurtner, "Perception of saturation in natural scenes," J. Opt. Soc. Am. A **33**, A194–A206 (2016).
57. M. Toscani and M. Valsecchi, "Lightness discrimination depends more on bright rather than shaded regions of three-dimensional objects," i-Perception **10**, 204166951988433 (2019).
58. M. Toscani, M. Valsecchi, and K. R. Gegenfurtner, "Optimal sampling of visual information for lightness judgments," Proc. Natl. Acad. Sci. USA **110**, 11163–11168 (2013).
59. M. Toscani, S. Zdravkovic, and K. R. Gegenfurtner, "Lightness perception for surfaces moving through different illumination levels," J. Vis. **16**(15), 21 (2016).
60. M. Toscani, M. Valsecchi, and K. R. Gegenfurtner, "Lightness perception for matte and glossy complex shapes," Vis. Res. **131**, 82–95 (2017).
61. M. G. Bloj, D. Kersten, and A. C. Hurlbert, "Perception of three-dimensional shape influences colour perception through mutual illumination," Nature **402**, 877–879 (1999).
62. A. I. Ruppertsberg and M. Bloj, "Reflecting on a room of one reflectance," J. Vis. **7**(13), 12 (2007).
63. C. Yu, M. Wijntjes, E. Eisemann, *et al.*, "Effects of inter-reflections on the correlated colour temperature and colour rendition of the light field," Lighting Res Technol. **55**, 772–793 (2022).
64. To represent the two hue directions in a more meaningful way than "clockwise" and "counterclockwise" shifts, we recoded them as "L-M-increasing" and "L-M-decreasing" shifts. Note that using *direction* as a factor in our model is incorrect from a theoretical standpoint, since it groups, for example, positive chroma shifts and L-M-increasing hue shifts together. These two levels are rather subsets of our chroma and hue shifts, respectively. We group them together for model simplicity but qualitatively assess them separately in *post hoc* analyses here.
65. L. Hedjar, M. Toscani, and K. R. Gegenfurtner, "Dataset: Importance of hue: color discrimination of three-dimensional objects and two-dimensional discs," Zenodo (2025), https://doi.org/10.5281/zenodo.13913190.