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

Perception of saturation in natural objects

LAYSA HEDJAR,[1,*] MATTEO TOSCANI,[2] AND KARL R. GEGENFURTNER[1]

[1]Abteilung Allgemeine Psychologie, Justus-Liebig-Universität-Gießen, 35394 Gießen, Germany
[2]Department of Psychology, Bournemouth University, Poole BH12 5BB, UK
*laysa.hedjar@psychol.uni-giessen.de

Received 29 September 2022; revised 2 January 2023; accepted 2 February 2023; posted 6 February 2023; published 23 February 2023

The distribution of colors across a surface depends on the interaction between its surface properties, its shape, and the lighting environment. Shading, chroma, and lightness are positively correlated: points on the object that have high luminance also have high chroma. Saturation, typically defined as the ratio of chroma to lightness, is therefore relatively constant across an object. Here we explored to what extent this relationship affects perceived saturation of an object. Using images of hyperspectral fruit and rendered matte objects, we manipulated the lightness–chroma correlation (positive or negative) and asked observers which of two objects appeared more saturated. Despite the negative-correlation stimulus having greater mean and maximum chroma, lightness, and saturation than the positive, observers overwhelmingly chose the positive as more saturated. This suggests that simple colorimetric statistics do not accurately represent perceived saturation of objects—observers likely base their judgments on interpretations about the cause of the color distribution.

https://doi.org/10.1364/JOSAA.476874

1. INTRODUCTION

The trichromaticity of human color vision is one of the most well-established results in vision science [1–6]. From this follows that color spaces can be broken down into three dimensions. These dimensions are relatively arbitrary, but the basic tenet is that each visual stimulus maps to a single point in color space. This is indeed the case for flat, matte surfaces or patches that uniformly emit light. However, the world around us consists mainly of three-dimensional objects, and the light reflected into the eye from these objects is affected by numerous factors such as shading or variations in pigmentation and material. The result is that the light entering the eye from an object forms a whole distribution of points in color space [7–10]. It is a major question how observers arrive at single estimates of hue, lightness, and saturation when being faced with these distributions. Do they simply take the mean of the corresponding values across all pixels?

Many objects actually have very little variation in hue [7,11,12] (but see [13]). However, lightness and chroma, as defined by CIELAB L^* and C^*, vary systematically. For many classes of surfaces, reflected light is derived from two mechanisms: a "specular" component which arises at the interface of the object surface and reflects the illuminant, and a "body" component which arises from the interaction of the illuminant and the diffuse object material and thus is informative about the spectral reflectance of the object [14–19]. Due to shading, the light from the body component forms a line in color space, passing through the darkest point in the color space and through the color coordinate of a surface point that is oriented perpendicular to the light source. This follows from Lambert's cosine law and results in a positive correlation between chroma and lightness. The law states that the intensity of the reflected light at a certain point on an object of a diffuse material is a function of the angle between the surface normal and the direction of the illuminant at that point. The overall incident light is multiplied by a factor of the cosine of that angle. If we apply this factor α, which decreases from 1 to 0 as the angle increases, to a reflectance spectrum and convert to color matching functions, e.g., CIE XYZ, the color coordinates are also simply multiplied by that factor. When converting XYZ to CIELAB L^* and C^*, the non-linear transformation changes the constant factor to $\sqrt[3]{\alpha}$ and the relationship between L^* and C^* is linear (for non-small XYZ values). Additional factors influencing the relationship between C^* and L^* can be, for example, interreflections between different parts of a scene, as could arise from a concavity within a colored object or from two colored objects near each other [20]. In this case, the lightness will decrease while the chroma increases.

Here, we investigate how human observers arrive at estimates of saturation for objects that are characterized by different distributions in color space. Are observers capable of disentangling different causes that lead to these distributions? Or do they simply take the average value of the color appearance metric "saturation" (C^*/L^*) across a distribution? Previously, we showed that the regions closest to the surface normal, least affected by shading, are mainly used as the basis for lightness judgments of objects [21,22]. Very little research, however, has

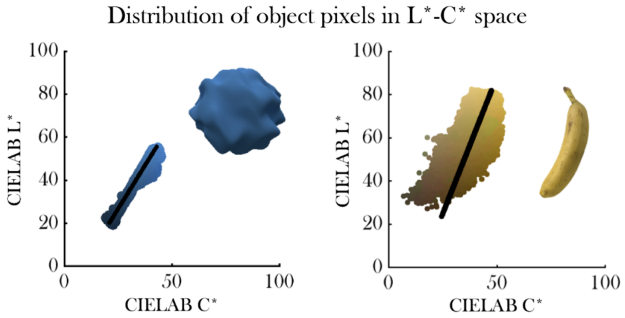

Fig. 1. Comparing variability in saturation measurements across objects. Both panels plot the distribution of pixels across an object in the L*-C* plane of CIELAB. The left panel plots the distribution of one rendered matte object, the right panel of a banana measured with a hyperspectral camera [7]. Depictions of the objects are inset on the right of each plot. The L*-C* distribution projected onto the first principal component (which explains 85% and 93% of the variance for the left and right plot, respectively) is plotted in black. Both plots show the positive correlation between lightness (L*) and chroma (C*) and demonstrate that variation due to shading minimally affects the chroma-to-lightness ratio.

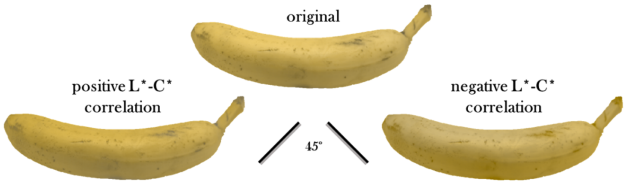

Fig. 2. Demonstration of the effect of L*-C* slope on the color appearance of an object. The pixels from a hyperspectral image of a banana were plotted in the L*-C* plane of CIELAB. The original banana (top) has a positive L*-C* slope at an angle of 68° (first principal component). For demonstration purposes, we rotated the distribution to 45° to amplify the appearance difference (bottom left image). The bottom right image shows the 45° banana with its distribution horizontally flipped to create a negative L*-C* slope. One can perceive a difference in object color between the positive-slope banana and the negative-slope banana. For the experiment, we performed additional manipulations to control for various colorimetric statistics.

been concerned with saturation judgments (but see [23–25]). Saturation is defined in most color appearance models as the ratio of chroma to lightness [24,26,27] and, given the above systematic relationship between lightness and chroma, this is logical. Figure 1 depicts the lightness–chroma relationship for a rendered matte object and a real object. The first principal component of each distribution (black line) explains over 85% of the variance and illustrates how shading has little to no effect on the chroma-to-lightness ratio.

We wanted to explore whether manipulations of this positive correlation between lightness and chroma affect saturation perception, holding all other variables constant. Figure 2 shows a demonstration of a banana whose L*-C* slope is inverted; one can see that the impression of the banana's color is different between the positive-slope and negative-slope version. While the banana with the positive C*/L* slope seems quite saturated, the banana with the negative slope appears much more pale and desaturated. This demonstration seems to indicate that the bananas with different C*/L* slopes are interpreted in different ways, even though their average C*/L* is approximately the same.

For our experiment, we created a more controlled set of stimuli with equivalent basic colorimetric statistics (C*, L*) but different lightness–chroma relationships. This was done by artificially manipulating natural and natural-looking stimuli which initially all have positive lightness–chroma correlations. We flipped the distribution of pixels of each stimulus on the L*-C* plane in CIELAB so that mean, median, and max C* and L* are the same, but the relationship between L* and C* is now negative: points with high lightness have low chroma and vice versa. Using both natural stimuli (fruits) and rendered "blobs," we show that observers almost always chose the stimulus with the positive L*-C* slope as having higher saturation than the stimulus with the negative L*-C* slope. Importantly, any simple calculation based on the standard definition (C*/L*) would

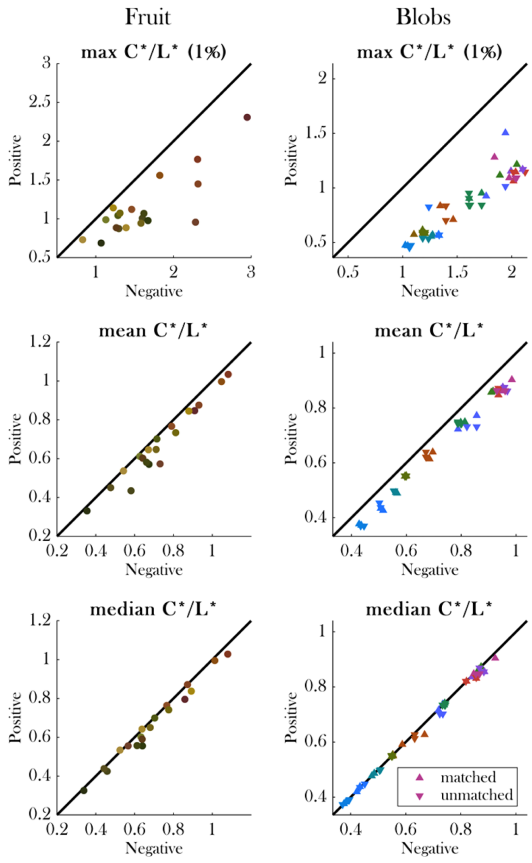

Fig. 3. Comparisons of saturation statistics between positive and negative object pairs. We compared max (top 1%), mean, and median saturation between fruit pairs and between blob pairs (matched- and unmatched-rotation). Each datapoint represents a pair, color-coded by the mean color of the object. For the blob plots (right), matched and unmatched blobs are plotted as upward- and downward-pointing triangles, respectively. Values for the negative object are plotted on the abscissa and for the positive object on the ordinate. Datapoints below the diagonal line indicate that the saturation statistic for the negative object was higher than for the positive object, and vice versa. For all pairs, datapoints lie below (or along) the diagonal line (tolerance of 0.005). C* and L* statistics (mean, median, max 1%) were equivalent between fruit pairs and matched blob pairs; for unmatched blob pairs, the negative-slope blob had higher values than the positive-slope.

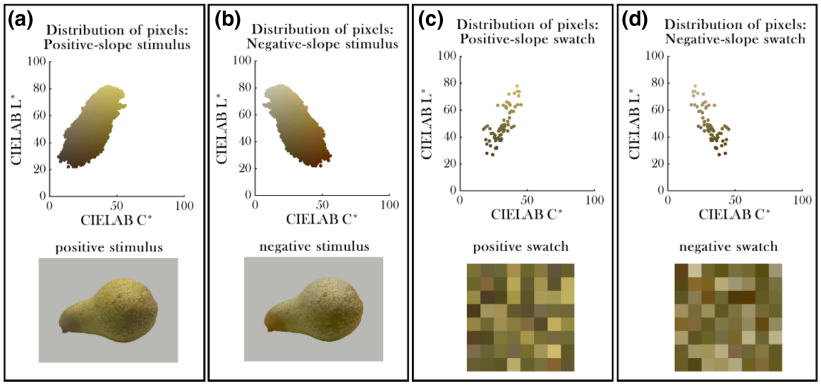

Fig. 4. Example of the pear stimulus pair used in the experiments, shown alongside its distributions in the L*-C* plane. The following are depicted: the distribution of the positive and negative pear [(a), (b), top] after symmetrizing and fixing out-of-gamut pixels, along with their corresponding images [(a), (b), bottom], and the distribution of pixels from the positive and negative swatches and the swatches themselves (c) and (d).

either yield equivalent predictions of saturation for each stimulus pair or would predict the negative-slope stimulus as more saturated.

Our results emphasize that simple colorimetric statistics calculated across object pixel distributions poorly predict saturation judgments. Rather, the human visual system seems to estimate the underlying causes of particular distributions of lightness and chroma. In the case of positive correlations, these are due to shading. In the case of negative correlations, the object is more likely to be achromatic, with the negative correlation stemming from lighting and interreflections.

2. METHODS

In the demonstrations above (Fig. 2), a few pixels fall out of the gamut of the monitor when the slope of the L*-C* line was adjusted. Therefore, we devised stringent methods to make sure that our experimental results were not caused by these underflows or overflows.

A. Stimuli

1. Fruit

Hyperspectral images of 19 fruit taken from a large database [7] were converted to $CIELAB_{LCH}$ color space. All 19 fruit distributions had positive L*-C* slopes (based on the first principal component). The hues of pixels whose C* value was less than 10 were set to the mean hue of the distribution; at such low chroma values, those pixels are almost achromatic and their given hue angles are likely to be due to noise. The body component which the positive L*-C* correlation is derived from is most informative about the spectral reflectance of an object; therefore, we avoided the rare fruits with high specularity or a negative native L*-C* correlation. In the natural world, most objects are matte; highly specular objects are unusual. In general, Tominaga and Wandell [19] found that the standard reflectance model consisting of a body component and a specular component can be used to describe the colors of fruits and plastics [14,17].

For each fruit stimulus, we "symmetrized" the C* values such that they were symmetrically distributed around the mean,

Fig. 5. Two of the fruit stimulus pairs (artichoke, papaya) used in the experiment. For each pair, the positive-slope stimulus is on the left and the negative-slope stimulus is on the right. Other fruit used in the experiment (not shown): nectarine, peach, dragonfruit, grapefruit, green apple, kiwi, lemon, lime, banana, carrot, mango, cherimoya, potato, pear, pomegranate, cucumber, zucchini (see [7,28]).

resulting in identical values for the C* mean and C* median. This was done by taking all C* values below (or above, depending on which side had the higher range of values) the median C*, computing their distances to the median, and applying those distances to the C* values of the points above (or below) the median; thus, for every point x C* units below the median C*, there was a point x C* units above the median C*. To create the negative-sloped L*-C* distributions, we horizontally flipped the C* values around the mean, keeping all else the same. In conjunction with the symmetrizing and flipping, we also identified pixels with values outside of the RGB gamut and shifted their values closer to the mean in L*-C* space until they were no longer out of range. Altogether, this resulted in negative and positive stimuli having exactly the same L* and C* statistics (max, mean, median, etc.) per fruit. We also calculated key statistics for saturation (C*/L*) and found that, across all fruit, the negative stimulus always resulted in higher or equal saturation values compared to the positive stimulus (tolerance = 0.005) (see Fig. 3). Figures 4(a) and 4(b) show a stimulus pair used in the experiment as well as its distributions plotted in the

L*-C* plane. Figure 5 depicts two of the 19 fruit pairs used in the experiment; the relative size of the stimuli is maintained.

2. Rendered Blobs

The advantage of using rendered objects is that we have greater control over the properties of the stimuli which can be varied accordingly. This also results in much less noise in the stimuli. Therefore, we rendered 3D lumpy matte "blobs" (Mitsuba 0.6) with different spatial rotations at 12 different hues in a diffuse lighting environment. We chose to render matte, non-specular stimuli to explicitly isolate the diffuse "body" component of the reflected light, which has a positive L*-C* correlation. The lighting environment was composed from high dynamic range images taken of the courtyard of the Doge's palace in Venice, Italy [29]. The scene is outdoors on a cloudy day. The matte and specular components of the blob material were defined using the Ward BRDF model [30,31] implemented in Mitsuba. The specular reflectance components were set to 0. The RGBs of the diffuse components were defined in $CIELAB_{LCH}$ with an L* of 70 and a C* of 50, with the white point defined as the white of the monitor (CIE1931 xyY: 0.3328, 0.3343, 142.35). The hue radians used ranged from 0.5 (28.6°) to 6.0 (343.8°) at steps of 0.5. We translated the distribution of pixels in L*-C* space as needed in order to minimize out-of-range values when flipping. In the same manner as with the fruit, we symmetrized the pixel values along the C* dimension; to create the negative-sloped stimulus, we horizontally flipped the C* values about the mean. This again resulted in equal C* and L* statistics; calculations of saturation (C*/L*) indicated that the negative-sloped stimulus across all variations had higher or equal saturation (tolerance = 0.005) (see Fig. 3). Figures 6(a) and 6(b) show the stimulus pair derived from a reddish blob along with its L*-C* distributions. For the experiment, we used two positive–negative pairs per hue. We called these "matched-rotation pairs" because the positive and negative stimulus within a pair had the same spatial rotation.

We also wanted to compare stimuli with slightly different spatial arrangements, in order to reduce the likelihood that observers focus on the same diagnostic region of the negative- and positive-sloped stimuli to make their judgments. Thus, we compared the statistics of the negative and positive stimuli across rotation sets (i.e., blobs with different spatial rotations) and gathered all pairs for which the negative stimulus of one rotation had higher L* and C* statistics than the positive stimulus of another rotation—"unmatched-rotation pairs." Because there was a limited number of unmatched-rotation pairs for which a negative blob had higher L* and C* statistics than a positive blob, the number used for each hue was not balanced (ranging from 0 to 4). We used a total of 24 matched-rotation stimulus pairs and 24 unmatched-rotation pairs. Figure 7 depicts a sample of matched- and unmatched-rotation blob pairs used in the experiment.

3. Swatches

The swatch stimuli served as a control in which the L*-C* slopes of the distributions were the same as those of the realistic stimuli but lacked their physical structure. For each stimulus (fruit and blob), we randomly sampled 64 pixels from the distribution and arranged them in an 8 × 8 matrix. The pixels chosen from the fruit and the matched-rotation blobs were matched with their corresponding pair [i.e., swatches from negative- and positive-slope stimulus pairs had the same C* and L* statistics, and C*/L* statistics were always higher (or equivalent) for the negative stimulus]. Figures 4(c) and 4(d) depict a swatch pair and its distributions derived from the pear and Figs. 6(c) and 6(d) from one matched-rotation blob. We created the unmatched swatch pairs in the same manner by which we created the unmatched-rotation blob pairs: comparing statistics of negative and positive swatches across blob rotation sets and gathering pairs for which a negative swatch had higher statistics than a positive swatch. The number of unmatched swatch pairs also was not balanced across hues and ranged from one to three pairs per hue. Nineteen swatch pairs were derived from the fruit (one from each fruit); 24 matched-rotation swatch pairs and 24 unmatched-rotation swatch pairs were derived from the blob sets.

B. Participants

Ten naïve observers (ages ranging from 20 to 29; median = 22.5) completed the paired comparisons task for the fruit and 10 different naïve observers (ages ranging from 19 to 35; median = 23.5) completed the task for the rendered blobs. All observers

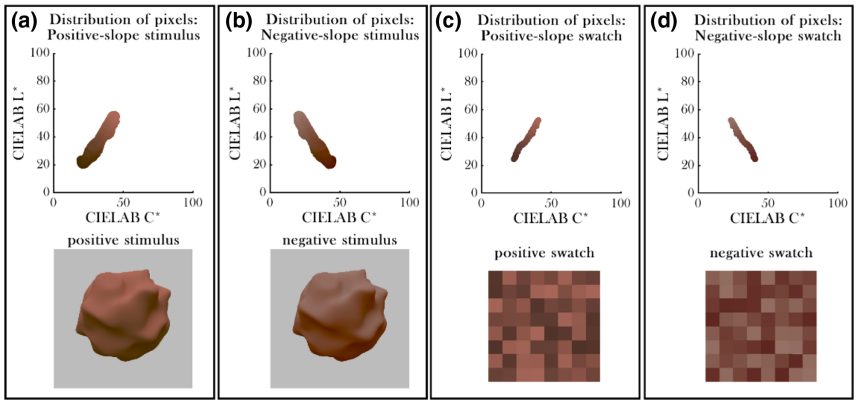

Fig. 6. Example of a matched-rotation blob stimulus pair used in the experiments, shown alongside its distributions in the L*-C* plane. The arrangement is the same as in Fig. 4.

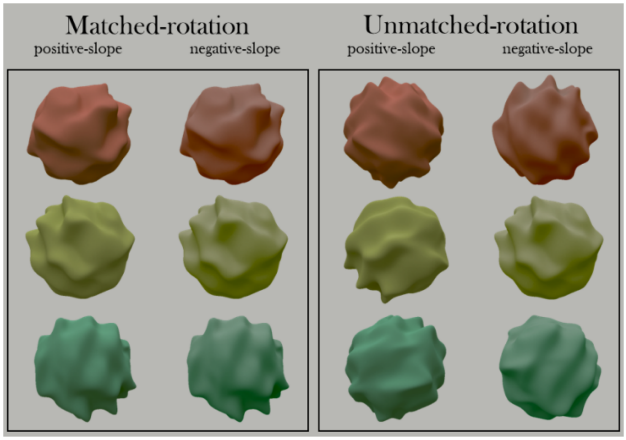

Fig. 7. Selection of blob pairs used in the experiment. Matched-rotation blob pair examples are shown on the lefthand side of the figure, unmatched-rotation pairs on the right. For each pair, the positive-slope stimulus is on the left and the negative-slope stimulus is on the right. Three of 12 possible hues are shown (rows).

gave informed consent and had normal color vision as assessed by the Ishihara Color Vision Test.

C. Procedure

Observers were asked to choose which of two objects presented side by side on the screen appeared more saturated. To orient their definition of saturation, we presented a green pepper at four levels of increasing saturation (generated by keeping L* constant and increasing C*) on the instructions screen.

Positive and negative slope pairs were displayed for one second on a mid-gray background (CIE1931 xyY: 0.3328, 0.3343, 71.1762) using an Eizo ColorEdge CG245W monitor with a resolution of 1920 × 1200 (50° × 32°). Observers sat 56 cm from the screen in a dark room. Fruit stimuli ranged in visual angle from about 7.76° to 30° in width and 6.41° to 18.7° in height. Blob stimuli spanned on average 14.3° × 14.3°. Swatches were all 13.9° × 13.9°. The stimuli were centered on the left and right halves of the screen (distance between centers of stimuli = 26°). After a pair of images was presented, observers indicated using the keyboard which image appeared more saturated. Observers responded at their own pace. Each fruit pair was shown 10 times and each blob pair was shown five times. For both experiments, presentation of all pairs was randomized, left–right positions of each stimulus pair were randomized, and object and swatch pairs were interleaved. The position of the 64 pixels of each swatch was randomized for every trial. Each experiment was separated into two blocks, which allowed for a short break in between.

3. RESULTS

For each negative–positive pair, we calculated the proportion of trials in which observers chose the stimulus with the positive L*-C* slope as more saturated than the negative L*-C* slope.

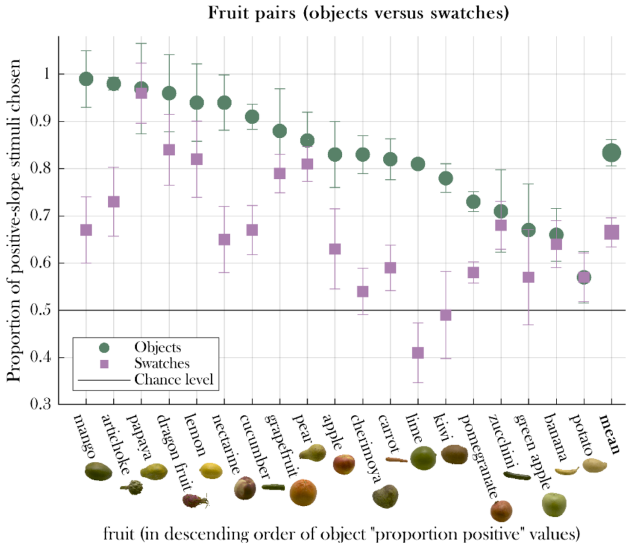

Fig. 8. Paired comparison task results separated by fruit. Each datapoint shows the proportion of responses in which the positive slope was chosen over the negative slope, averaged across observers. Swatch stimuli (purple squares) and object stimuli (green circles) are plotted for each fruit. The means across all fruit are plotted on the far right. A horizontal line at 0.5 indicates chance-level performance. Error bars indicate ±1 standard error of the mean (SEM).

A. Fruit

Figure 8 plots the results averaged across participants for the fruit. For all conditions except two (lime swatch and kiwi swatch), observers chose the positive stimulus more often than the negative stimulus as more saturated. This was reflected in a one-tailed sign test for both fruit and swatches (Bonferroni-corrected $p < 0.001$ for fruit and for swatches). We tested whether results for fruit and swatches were significantly different using a non-parametric Mann–Whitney U test with stimulus type as a factor. The mean proportion of positive-slope stimuli chosen for the fruits was significantly higher than for the swatches (0.83 versus 0.67; $Z = 3.40$, $p < 0.001$).

B. Rendered Blobs

Figure 9(a) plots the results for the rendered blobs and their swatches across hues. A one-tailed sign test showed that "proportion positive" results for blobs and their swatches were significantly greater than chance (Bonferroni-corrected $p = 0.002$ for blobs and for swatches). A Mann–Whitney U test with matched–unmatched as a factor indicated no significant difference in the proportion of positive-slope stimuli chosen between matched and unmatched pairs (0.86 versus 0.82, respectively; $Z = 0.98$, $p = 0.33$). We did find a significant difference when using stimulus type as a factor: the proportion of positive-slope stimuli chosen for object pairs was higher than that for swatch pairs (0.93 versus 0.75; $Z = 2.99$, $p = 0.003$). We also explored possible differences between hues. Figure 9(b) plots the proportion of positive-slope stimuli chosen, averaged across all pairs of each stimulus type and separated by hue. A Kruskal–Wallis test for each stimulus type indicated that there was no significant difference between hues (objects: $H(11) = 9.31$; $p = 0.59$, swatches: $H(11) = 18.8$, $p = 0.06$).

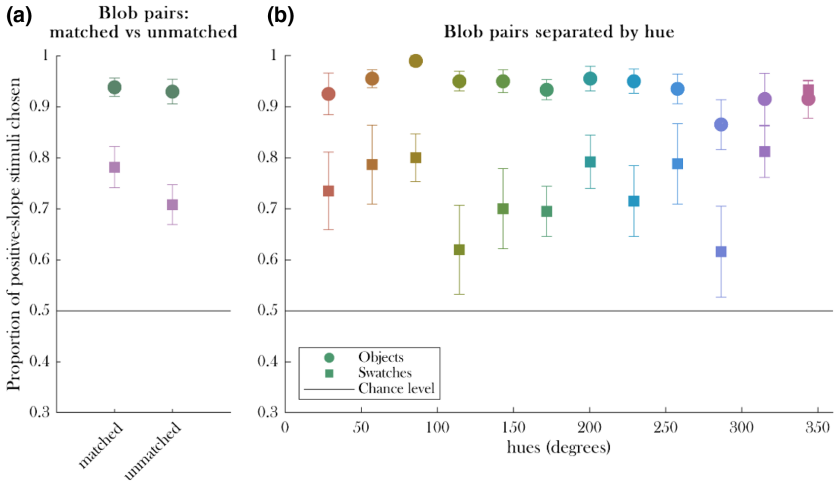

Fig. 9. Paired comparison task results for blob stimuli. The *y* axis plots the proportion of responses in which the positive slope was chosen over the negative slope. A horizontal line at 0.5 indicates chance-level performance. (a) Mean values separated for matched and unmatched blobs (green circles) and swatches (purple squares). Results are averaged across participants and hues. (b) Mean values separated by hue (averaged across participants and pair type). Error bars indicate ±1 SEM.

"Proportion positive" for the swatches shows more variability across hues than for the blob objects, with some swatch means being very close to object means (pinkish hue) while others much more different (yellowish hue). It is unclear what might drive this difference, but observers may be using very different strategies to make saturation judgements about the swatches.

4. DISCUSSION

We explored whether colorimetric measures of saturation are indicative of perceived saturation in stimuli which are defined by color distributions, such as realistic objects. Our findings suggest that judgments of saturation depend on the relationship between pixels in the L*-C* plane rather than a single estimate of saturation, chroma, or lightness computed across all pixels of the object. We demonstrate this with calibrated images of natural object stimuli and artificial computer graphics object stimuli (fruit and rendered "blobs," respectively). Notably, this effect arises even for object pairs of differing spatial structure, indicating that participants do not base their choice on the same physical point on both objects. Lastly, we find a weaker, but still significant, instance of the effect with colored mosaics (swatches), which have little ecological relevance.

A number of previous studies have looked at ensemble color perception of such multicolored mosaics. Kuriki [32] and Kimura [33] both presented observers with equiluminant mosaics that varied in hue and chroma. Observers' judgments were biased toward the most saturated patch when making global color judgments of the mosaic. Similarly, Sunaga and Yamashita [34] showed observers multicolored mosaics of equal hue and equal brightness but differing saturation and found that observers often gave more weight to the most saturated elements. Choi and colleagues [35] also found that observers' "single color impressions" of monochromatic stripes of low and high luminance were biased toward higher luminance and higher chroma values. Altogether these studies indicate that observers do not simply average across all present colors to form a global color impression of a patch. Their judgments are often biased toward the most saturated elements of the mosaics.

Our study supplements this work by showing that observers do not base their judgments solely on the most saturated pixels of an object or mosaic. We test this by pitting two stimuli against each other: both stimuli have equal lightness and equal chroma variation but inverse lightness and chroma relationships (one negatively correlated, one positively correlated). We chose stimulus pairs such that one in the pair presented with higher mean and max saturation, chroma, and lightness than the other. If observers based their judgments on the values of a single dimension (i.e., mean chroma, max saturation, etc.), they would have chosen the negative-slope stimulus as more saturated than the positive-slope stimulus, since the negative-slope stimulus had higher or equal mean/median/max values than the positive-slope stimulus. We instead find the reverse: observers consistently chose the positive-slope stimulus as more saturated. Importantly, we show that this effect is significantly greater for objects than for colored mosaics, suggesting that observers develop some interpretation of the scene and consider its influence on the pixel distribution.

Giesel and Gegenfurtner [11] explored representative color judgments of objects by having observers make color matches of real-world objects to uniform patches presented on a screen. The objects were of various colors and materials, such as wool, paper, and candlewax. They found that observers tended to overshoot chroma and lightness judgments in comparison to the means of the objects. Similarly, our observers do not use mean chroma, lightness, or saturation to make judgments of our stimuli. Interestingly in Giesel and Gegenfurtner's study, observers overshot their chroma judgments to a greater extent for objects with lower mean lightness than lighter objects, suggesting at least that the observers take lightness into account when making chroma judgments. However, from their study and ours, it is still unclear what diagnostic observers use to make judgments of chroma or saturation.

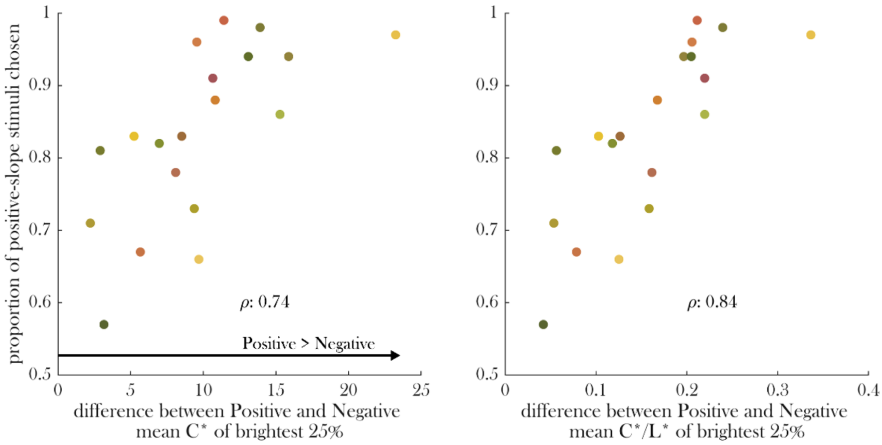

Fig. 10. Plots comparing the mean C* and C*/L* values of the 25% brightest pixels between positive–negative fruit pairs. The abscissa plots the difference between positive and negative pairs such that a positive value indicates that the positive pair had a larger statistic than the negative. The ordinate plots the proportion of trials (averaged across observers) for which the positive-slope stimulus was chosen. Pairs are color-coded by their mean stimulus color. We computed Spearman's rho for each plot. Correlations were relatively high for both statistics (0.74 for mean C*, 0.84 for mean C*/L*; $p < 0.001$). For swatches (not shown), correlations were weaker but still positive (0.51 and 0.48, respectively; $p < 0.05$).

While hue tends to have little variation within matte objects of a single reflectance, chroma and lightness, and therefore saturation, can have much variation depending on the interaction between the direction of illumination and the surface normal across the object (see Fig. 1). Observers may use a diagnostic region of the object (or of the L*-C* curve) to judge saturation, as they do with lightness. We know that estimates of lightness of objects are based on "brighter, diffusely shaded regions," which are particularly diagnostic for their achromatic reflectance, and that observers tend to fixate on these regions when asked to make lightness judgments [21,22,36]. The region of an object diagnostic for its saturation may depend on interpretations of illumination and interreflections within the scene. We explored whether observers might be using two dimensions simultaneously from which to derive their saturation judgments. Figure 10 plots the C* and C*/L* of the brightest pixels for the fruit. Blob comparisons (not shown) are similar but have less variation, overlaying the fruit points in the region of high "proportion positive" results (correlation coefficients were insignificant). On the abscissa is the difference between the positive and negative stimulus pair for that statistic; positive values indicate that the positive stimulus has a higher value. On the ordinate are plotted the average "proportion positive" values across observers. Note that for these statistics, the positive pair always has higher mean C* and C*/L* values for the brightest 25% pixels. Spearman's correlation coefficients indicate high correlations ($\rho = 0.74$ and 0.84 for C* and C*/L* of brightest 25%, respectively): the greater the difference between a pair, the more likely the observers choose the positive pair. If observers are heavily biased toward the brighter pixels of the object, our data suggest that their saturation judgments are driven by the differences in saturation or chroma of these brighter regions. These regions may be more informative and less noise-prone than the rest of the object or mosaic. However, a more carefully controlled experiment is needed to make this conclusion.

In our study, we use realistic objects whose native L*-C* slope is positive. We artificially invert this relationship to create objects whose L*-C* slope is negative. Are there real-world scenarios in which a negative L*-C* correlation would arise? Do observers perceive the negative-slope stimuli as if in these scenarios? Figure 11 displays simulations of two such scenarios. In the first [Fig. 11(a)], the matte object and the illumination are white, but the (matte) walls are chromatic. The second example [Fig. 11(b)] depicts a translucent stimulus on an achromatic background sitting in a diffuse lighting environment. The distribution of pixels emitted from both sets of stimuli have negative L*-C* correlations. Observers may have interpreted our negative-slope stimuli under one of these two scenarios, for which the reflectance characteristics of the stimuli are quite desaturated (a matte white material and a translucent colored material). If this is their interpretation of the scene, then their perception of the negative-slope stimuli as less saturated would fit. The observer may attribute different regions of the L*-C* correlation line to different effects of lighting and/or interreflections—for example, in the scenario on the left, the

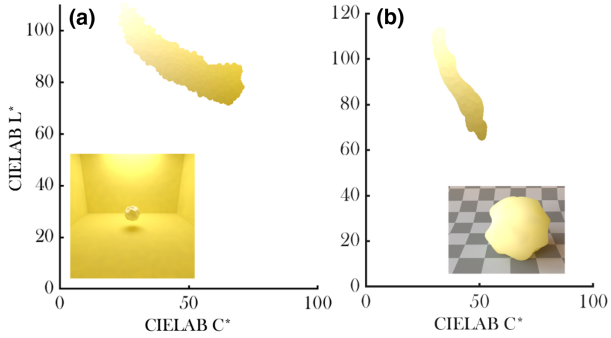

Fig. 11. Rendered scenarios in which the distribution of pixels reflected from an object has a negative slope when plotted on the L*-C* plane of CIELAB space. (a) A scenario in which a matte object has a white reflectance and the illuminant emits white light (RGB = [1,1,1] for both), but the matte walls have a yellowish reflectance. (b) A scene of a yellow translucent object rendered on an achromatic checkered background under a diffuse lighting environment.

brightest and least chromatic pixels mostly come from the upper surface of the blob, which reflects the greatest proportion of direct light from the (white) illuminant, while the darkest and most chromatic pixels result largely from the interreflections between the lower surface of the blob and the (yellow) floor. But merely exploiting the polarity of the L*-C* correlation to judge saturation is likely an oversimplification. Additionally, we know that observers are sensitive to the "naturalness" of stimuli, which can impact perception of saturation and colorfulness [37], and we cannot be sure that our observers perceived our artificial and manipulated stimuli as realistic. Objects with more complex material properties may also elicit a negative L*-C* correlation, and potential heuristics used for matte objects likely do not transfer as well to glossy [38] or translucent objects, or to specialized subcategories of objects such as faces [39,40].

Overall, we showed here that the standard definition of saturation (C*/L*) does not accurately reflect saturation judgments of realistic objects and colored mosaics. While it is still unclear what computations observers use to make judgments about saturation, we suggest that they make some interpretation of the scene and the sources for the reflected light, considering the relationship between dimensions of color space, and then give more weight to pixels which are more diagnostic of the object's color, such as the brighter pixels. Our results imply that color perception of real-world stimuli in complex scenes leads to judgements of object color that are not best represented by simple colorimetric calculations. Rather, the visual system likely makes interpretations about the environment and uses this to determine to what extent the color variation can be attributed to invariant properties of the object.

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

Acknowledgment. We are grateful to Eli Brenner, Cehao Yu, and Shin'ya Nishida for helpful comments on previous presentations of this work. We would also like to thank Bianca Baltaretu for editorial assistance.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper, relevant statistics of the stimuli, and pngs of the stimuli used are available in Ref. [41]. The hyperspectral image set from which the present fruit stimuli were derived can be found at [28] (see Ref. [7]).

REFERENCES

1. H. Grassmann, "Zur Theorie der Farbenmischung," Ann. Phys. **165**, 69–84 (1853).
2. H. Helmholtz, *Handbuch der Physiologischen Optik*, Vol. 9 of Allgemeine Encyklopädie der Physik (Voss, 1867).
3. D. H. Krantz, "Color measurement and color theory: I. Representation theorem for Grassmann structures," J. Math. Psychol. **12** (3), 283–303 (1975).
4. J. C. Maxwell, "1. Experiments on colour as perceived by the eye, with remarks on colour-blindness," Proc. R. Soc. Edinburgh B **3**, 299–301 (1857).
5. J. L. Schnapf, T. W. Kraft, and D. A. Baylor, "Spectral sensitivity of human cone photoreceptors," Nature **325**, 439–441 (1987).
6. A. Stockman and L. T. Sharpe, "Cone spectral sensitivities and color matching," in *Color Vision: From Genes to Perception* (Cambridge University Press, 1999), pp. 53–88 .
7. R. Ennis, F. Schiller, M. Toscani, and K. R. Gegenfurtner, "Hyperspectral database of fruits and vegetables," J. Opt. Soc. Am. A **35**, B256–B266 (2018).
8. T. Hansen and K. R. Gegenfurtner, "Color scaling of discs and natural objects at different luminance levels," Vis. Neurosci. **23**, 603–610 (2006).
9. M. Vurro, Y. Ling, and A. C. Hurlbert, "Memory color of natural familiar objects: Effects of surface texture and 3-D shape," J. Vis. **13**(7): 20 (2013).
10. C. Witzel and K. R. Gegenfurtner, "Color perception: objects, constancy, and categories," Annu. Rev. Vis. Sci. **4**(1), 475–499 (2018).
11. M. Giesel and K. R. Gegenfurtner, "Color appearance of real objects varying in material, hue, and shape," J. Vis. **10**(9): 10 (2010).
12. C. Witzel and H. Dewis, "Why bananas look yellow: The dominant hue of object colours," Vision Res. **200**, 108078 (2022).
13. Z. Milojevic, R. Ennis, M. Toscani, and K. R. Gegenfurtner, "Categorizing natural color distributions," Vision Res. **151**, 18–30 (2018).
14. G. J. Klinker, S. A. Shafer, and T. Kanade, "A physical approach to color image understanding," Int. J. Comput. Vis. **4**(1), 7–38 (1990).
15. J. Mollon, "Monge: The Verriest Lecture, Lyon, July 2005," Vis. Neurosci. **23**, 297–309 (2006).
16. G. Monge, "Mémoire sur quelques phénomènes de la vision," Ann. Chim. **3**, 131–147 (1789).
17. S. A. Shafer, "Using color to separate reflection components," Color Res. Appl. **10**(4), 210–218 (1985).
18. S. Tominaga, "Surface identification using the dichromatic reflection model," IEEE Trans. Pattern Anal. Mach. Intell. **13**(7), 658–670 (1991).
19. S. Tominaga and B. A. Wandell, "Standard surface-reflectance model and illuminant estimation," J. Opt. Soc. Am. A **6**, 576–584 (1989).
20. C. Yu, M. Wijntjes, E. Eisemann, and S. Pont, "Effects of interreflections on the correlated colour temperature and colour rendition of the light field," Light. Res. Technol. (to be published).
21. 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).
22. M. Toscani, M. Valsecchi, and K. R. Gegenfurtner, "Selection of visual information for lightness judgements by eye movements," Philos. Trans. R. Soc. B **368**, 20130056 (2013).
23. R. Cao, M. Castle, W. Sawatwarakul, M. Fairchild, R. Kuehni, and R. Shamey, "Scaling perceived saturation," J. Opt. Soc. Am. A **31**, 1773–1781 (2014).
24. F. Schiller, M. Valsecchi, and K. R. Gegenfurtner, "An evaluation of different measures of color saturation," Vision Res. **151**, 117–134 (2018).
25. F. Schiller and K. R. Gegenfurtner, "Perception of saturation in natural scenes," J. Opt. Soc. Am. A **33**, A194–A206 (2016).
26. M. D. Fairchild, *Color Appearance Models*, 3rd ed. (Wiley, 2013).
27. M. R. Luo and C. Li, "CIECAM02 and its recent developments," in *Advanced Color Image Processing and Analysis*, C. Fernandez-Maloigne, ed. (Springer, 2013), pp. 19–58.
28. https://www.allpsych.uni-giessen.de/GHIFVD.
29. P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *ACM SIGGRAPH 2008 Classes (SIGGRAPH)* (Association for Computing Machinery, 2008), pp. 1–10.
30. D. Geisler-Moroder and A. Dür, "A new ward BRDF model with bounded Albedo," Comput. Graph. Forum **29**, 1391–1398 (2010).
31. G. J. Ward, "Measuring and modeling anisotropic reflection," Comput. Graph. **26**, 265–272 (1992).
32. I. Kuriki, "Testing the possibility of average-color perception from multi-colored patterns," Opt. Rev. **11**, 249–257 (2004).
33. E. Kimura, "Averaging colors of multicolor mosaics," J. Opt. Soc. Am. A **35**, B43–B54 (2018).
34. S. Sunaga and Y. Yamashita, "Global color impressions of multicolored textured patterns with equal unique hue elements," Color Res. Appl. **32**(4), 267–277 (2007).
35. S. H. Choi, H. Kim, K.-C. Shin, H. Kim, and J.-K. Song, "Perceived color impression for spatially mixed colors," J. Disp. Technol. **10**(4), 282–287 (2014).

36. M. Toscani and M. Valsecchi, "Lightness discrimination depends more on bright rather than shaded regions of three-dimensional objects," i-Perception **10**(6), 204166951988433 (2019).
37. T. Masumitsu and Y. Mizokami, "Influence of naturalness of chroma and lightness contrast modulation on colorfulness adaptation in natural images," J. Opt. Soc. Am. A **37**, A294–A304 (2020).
38. M. Toscani, M. Valsecchi, and K. R. Gegenfurtner, "Lightness perception for matte and glossy complex shapes," Vision Res. **131**, 82–95 (2017).
39. M. Hasantash, R. Lafer-Sousa, A. Afraz, and B. R. Conway, "Paradoxical impact of memory on color appearance of faces," Nat. Commun. **10**(1), 3010 (2019).
40. H. Yoshikawa, K. Kikuchi, H. Yaguchi, Y. Mizokami, and S. Takata, "Effect of chromatic components on facial skin whiteness," Color Res. Appl. **37** (4), 281–291 (2012).
41. L. Hedjar, M. Toscani, and K. R. Gegenfurtner, "Dataset and stimuli: perception of saturation in natural objects," Zenodo (2023), https://doi.org/10.5281/zenodo.7644301.