

# Geometrical structure of perceptual color space: Mental representations and adaptation invariance

Robert J. Ennis

Justus-Liebig Universität,  
Allgemeine Psychologie Abteilung,  
Gießen, Deutschland



Qasim Zaidi

State University of New York,  
Graduate Center for Vision Research,  
New York, NY, USA



Similarity between percepts and concepts is used to accomplish many everyday tasks, e.g., object identification; so this similarity is widely used to construct geometrical spaces that represent stimulus qualities, but the intrinsic validity of the geometry, i.e., whether similarity operations support a particular geometry, is almost never tested critically. We introduce an experimental approach for equating relative similarities by setting perceived midpoints between pairs of stimuli. Midpoint settings are used with Varignon's Theorem to test the intrinsic geometry of a representation space, and its mapping to a physical space of stimuli. For perceptual color space, we demonstrate that geometrical structure depends on the mental representation used in judging similarity: An affine geometry was valid when observers used an opponent-color mental representation. Similarities based on a conceptual space of complementary colors thus power a geometric coordinate system. An affine geometry implies that similarity can be judged within straight lines and across parallel lines, and its neural coding could involve ratios of responses. We show that this perceptual space is invariant to changes in illumination color, providing a formal justification to generalize color constancy results measured for color categories, to all of color space. The midpoint measurements deviate significantly from midpoints in the extensively used "uniform" color spaces CIELAB and CIELUV, showing that these spaces do not provide adequate metric representation of perceived colors. Our paradigm can thus test for intrinsic geometrical assumptions underlying the representation space for many perceptual modalities, and for the extrinsic perceptual geometry of the space of physical stimuli.

## Introduction

Sensory organs provide organisms with clues about the environment, but the properties relevant to drive

behavior are rarely explicit in the sensory input. To facilitate dealing with the environment, populations of neurons build and manipulate representations that make useful properties of materials, objects, illuminations, and atmospheres available to nonsensory processes. In this paper we examine fundamental properties of these representations, formulated as perceptual spaces. The methods we devise are of general applicability, but this paper concentrates on color perception.

A perceptual space consists of a set of relevant stimuli along with a set of similarity relationships. Perceptual spaces have been constructed for features as diverse as gloss (Ferwerda, Pellacini, & Greenberg, 2001; Wills, Agarwal, Kriegman, & Belongie, 2009), patterns (Victor & Conte, 2012), timbre (Lakatos, 2000; Terasawa, Slaney, & Berger, 2005), vowels (Pols, van der Kamp, & Plomp, 1969), sound textures (McDermott, Schemitsch, & Simoncelli, 2013), gestures (Arfib, Couturier, Kessous, & Verfaille, 2002), biological motion (Giese & Lappe, 2002), tactile textures (Hollins, Bensmaia, Karlof, & Young, 2000), tactile orientation (Bensmaia, Denchev, Dammann, Craig, & Hsiao, 2008), odors (Cleland, Johnson, Leon, & Linster, 2007), and others (Zaidi et al., 2013). The characteristics of every perceptual space center on two fundamental properties: dimensionality and intrinsic geometry, which are, in turn, consequences of the space's metric, i.e., the operation that defines similarity. Historically, similarities have been estimated by errors in matches as estimates of just-discriminable differences, or thresholds, or numerical ratings. Based on the experimentally determined properties of the similarity measure, the perceptual space can be assigned a well-defined geometry, thus providing access to a large number of theorems that in turn specify implications of the representational structure. These geometries form a

Citation: Ennis, R. J., & Zaidi, Q. (2019). Geometrical structure of perceptual color space: Mental representations and adaptation invariance. *Journal of Vision*, 19(12):1, 1–17, <https://doi.org/10.1167/19.12.1>.

<https://doi.org/10.1167/19.12.1>

Received March 15, 2019; published October 1, 2019

ISSN 1534-7362 Copyright 2019 The Authors



natural hierarchy, with more highly structured geometries placing greater demands on the conditions that the metric must satisfy (Klein, 1939; Brannan, Esplen, & Gray, 1999). At the top of the hierarchy is Euclidean geometry and its non-Euclidean relatives elliptical and hyperbolic, which allow representing stimuli as vectors, with sizes and angles invariant to transformations. Affine geometry is one step down the hierarchy: It allows for vector representations on arbitrarily scaled axes; thus, lines and parallelism remain invariant to transforms, but angle or size do not. Further down is Projective geometry: Collinearity and dimension remain invariant, but parallelism does not. Lower still, with the fewest geometrical requirements, is Topological space, where proximity is invariant but collinearity and dimension are not. Superimposed on this characterization of the intrinsic geometry of the perceptual space is its extrinsic geometry, which is the mapping of the perceptual space onto a physical space of stimuli, characterizing which can provide additional information about neural transformations. In terms of the dimensionality of the space, higher dimensional representations enable added flexibility in learning and finer grained qualitative distinctions, but can impose a higher cost on similarity computations.

Color spaces based on intuition and/or practical experience with mixing pigments date from antiquity (Smithson, Dinkova-Bruun, Gasper, Huxtable, McLeish, & Panti, 2012; Kirchner, 2015), but Maxwell (1860) provided a paradigmatic example of the empirical analysis of a stimulus space. By restricting empirical operations to color matches, irrespective of color percepts, and showing that color matches satisfy the linearity properties of additivity and scalar multiplication, Maxwell was able to embed color matches into the structure of a linear algebra, thus allowing for vector operations to predict the results of combining lights of different colors. Moreover, although the physical combinations of visible lights that range from 400–700 nm form a space of infinite dimensionality, Maxwell showed that color-matches fitted into a more tractable three-dimensional space. The key observation was that color-normal observers could match any light by adjusting the intensities of any three spectrally fixed lights, as long as none of the three primary lights was matchable by a combination of the other two. To probe the physiological basis of color space, Maxwell showed that the matches of congenital dichromats were a reduced subset of color-normal matches, so that the spectral sensitivities of the three types of cones could be obtained from a linear transform of the color matching functions using the confusion vectors of congenital dichromats. Electrophysiological measurements of the spectral sensitivities of human cones (Schnapf, Kraft, & Baylor, 1987) confirmed estimates using color matches supplemented by psychophysical measurements of

isolated cone sensitivities (König & Dieterici, 1886; Smith & Pokorny, 1975; Stockman & Sharpe, 2000). Maxwellian spaces based on metamers (physically distinct stimuli that appear identical), e.g., CIEXYZ1931 and Macleod-Boynton (MacLeod & Boynton, 1979), have proven invaluable for psychophysical and physiological investigations of the visual system. However, they only tell us that two physically distinct stimuli will look the same, and they do not tell us what color any stimulus will be. In formal terms, they do not provide a basis for representing similarity because their axes can be scaled arbitrarily without altering metamers, so that relative distances along nonparallel lines are incommensurable.

There have been many attempts to add additional structure to Maxwellian spaces to represent perceived distances between colors. The most systematic attempts have built on MacAdam's (1942) color ellipses. Using an extensive set of data collected on one observer, MacAdam found that errors in making color matches were roughly elliptical in shape and their orientation and size changed systematically in CIEXYZ space. Under the assumption that every ellipse represents a unit of perceived color difference, a transform of CIEXYZ axes that turns all ellipses into approximate circles of similar radius imposes an isotropic Euclidean metric. Theoretically motivated approximations were proposed by Le Grand (1949) and Friele (1961), and in some ways these representations are close to later revealed color properties of retinal ganglion cells (Sun, Smithson, Zaidi, & Lee, 2006) and LGN cells (Derrington, Krauskopf, & Lennie, 1984), but industry relies heavily on “uniform” color spaces such as CIELAB (Wyszecki & Stiles, 1982) and CIELUV (Wyszecki & Stiles, 1982), which use nonlinear transformations based on empirical criterion. All these spaces suffer from the limitation that each ellipse estimates just noticeable color differences from each local color, without accounting for the effects of local adaptation states around separated colors (Krauskopf & Gegenfurtner, 1992). “Uniform” color spaces are often used to choose sets of equally spaced colors spanning color space for psychophysics and electrophysiology experiments, but the validity of the equal spacing is uncertain because such sets have not been critically tested against psychophysically measured perceived similarities. We do this testing in the final section of this paper.

Since relative color similarities are regularly used to make inferences about the environment, and guide action, e.g., to identify materials and surfaces across spectrally distinct illuminations (Zaidi, 1998; Zaidi & Bostic, 2008; Radonjic, Cottaris, & Brainard, 2015), similarity operations have been widely used to provide geometrical structure to color spaces. Multidimensional scaling (Shepard, 1962) is the most common method to

specify color spaces based on numerical ratings of similarity between colors (Indow, 1980), with all the complications of mapping a perceptual quality to subjective numbers, and the Euclidean geometry may not be justified. Wuerger and colleagues (Wuerger, Maloney, & Krauskopf, 1995) showed that proximity judgments between colors fail Euclidean assumptions. Thus, the intrinsic geometry of color space remains to be determined: It should have enough structure to support judgments of relative similarity, but proximity judgments indicate that it is not Euclidean.

We introduce a method to directly investigate the geometrical structure supporting a color similarity space. Varignon's Theorem (Coxeter & Greitzer, 1967) states that the bimedians of a quadrilateral bisect each other, i.e., the point of intersection of the two straight lines joining the midpoints of opposite sides is the midpoint of both lines (Appendix Figure A1). This theorem holds only for geometrical spaces where stimuli can be represented as vectors, i.e., with an Affine or higher structure. If the vertices of the quadrilateral can be expressed as vectors, then the overlapping midpoints provide two different ways of estimating the same centroid vector (Appendix Figure A2). So, if Varignon's theorem does not hold, then the space is not Affine. We tested whether perceptual color space is affine by estimating pairs of colors set as centroids of a quadrilateral covering extended areas of color space and seeing if the two coincided in accordance with Varignon's Theorem. Observers viewed a test patch flanked by two patches, each containing one vertex color of the test quadrilateral. They were instructed that a midpoint between two colors is the color that is simultaneously most similar to the two, and could be ascertained by first finding the set of stimuli that are equally similar to the two fixed stimuli, and then from this set, the stimulus that is most similar to both. After finding the midpoints for the four sides, observers set the midpoints for each of the two pairs of facing midpoints. These midpoint settings were not close to each other for any observer in any condition, thus refuting the Affine assumption. Since color judgments based on "reddish-greenish" and "bluish-yellowish" opponent-dimensions give very stable estimates of color categories (Chichilnisky & Wandell, 1999; Smithson & Zaidi, 2004), observers were then instructed to consider the color difference between the endpoints along the opponent dimensions, and to adjust the middle patch's hue and saturation to a color perceived as the midpoint on both dimensions, i.e., equally and most similar to both endpoints. For all observers, the two final midpoints for each quadrilateral coincided, thus satisfying the conditions for an Affine space. Therefore, when observers explicitly use an opponent-color mental representation to judge color similarities, the perceptual color space of relative

similarities has an Affine structure. A Euclidean color space would enable the distance between any two colors to represent magnitude of similarity, and this is not possible in the weaker Affine space. However, in an Affine space, ratios of distances along every color line do provide measures of relative similarity, and parallelism does provide similarity between color changes.

If similarities are represented as neural responses in the brain, then to be widely useful, these responses must be invariant across conditions, just as some extrastriate neurons have object sensitivities that are invariant to pose (Pinto, Doukhan, DiCarlo, & Cox, 2009). We show that the geometrical space constructed with the midpoint settings is invariant across different overlaid colored illuminants, and this has significant implications for color constancy (Smithson & Zaidi, 2004; Zaidi & Bostic, 2008).

## Experiment 1: Geometrical test of structure for perceptual color space

We used Varignon's Theorem to test if perceptual color space is Affine, by translating it into a series of psychophysical midpoint judgments (see Supplemental Methods). After adapting to a midgray background, three rectangles appeared on a calibrated color monitor (Figure 1A). The colors of the two outside rectangles were set to two adjacent vertices of a quadrilateral in the MacLeod-Boynton (MacLeod & Boynton, 1979) equi-luminant color plane, and observers were instructed to use a joystick to set the color of the central rectangle to the perceptual midpoint of the end-point colors by finding the color that was the most similar to both end-point colors, out of all colors equally similar to both. This was repeated for the four sides defining each color quadrilateral. Then, the mean settings for opposite sides were used as endpoints, and observers were instructed to find the perceptual midpoint. The two midpoints for the two pairs of these endpoints are both estimates of the centroid vector, if the four vertices of the quadrilateral can be represented as vectors in an affine space. Therefore, if the final two midpoint settings did not coincide, an intrinsic Affine geometry was rejected for the perceptual color space. Means and standard deviations of midpoint settings were calculated for 10 repeated measurements by each of four observers, for a large square and a large diamond that were centered at the midgray background color and spanned most of the equiluminant color plane displayable on the monitor (Figure 1B: The endpoint pairs in the diamond are separated by diagonal lines, while those in the squares are separated by lines parallel to the S/L + M or L/L + M axes). For all observers, and all quadrilaterals, the two centroid-estimating midpoints did not come close to

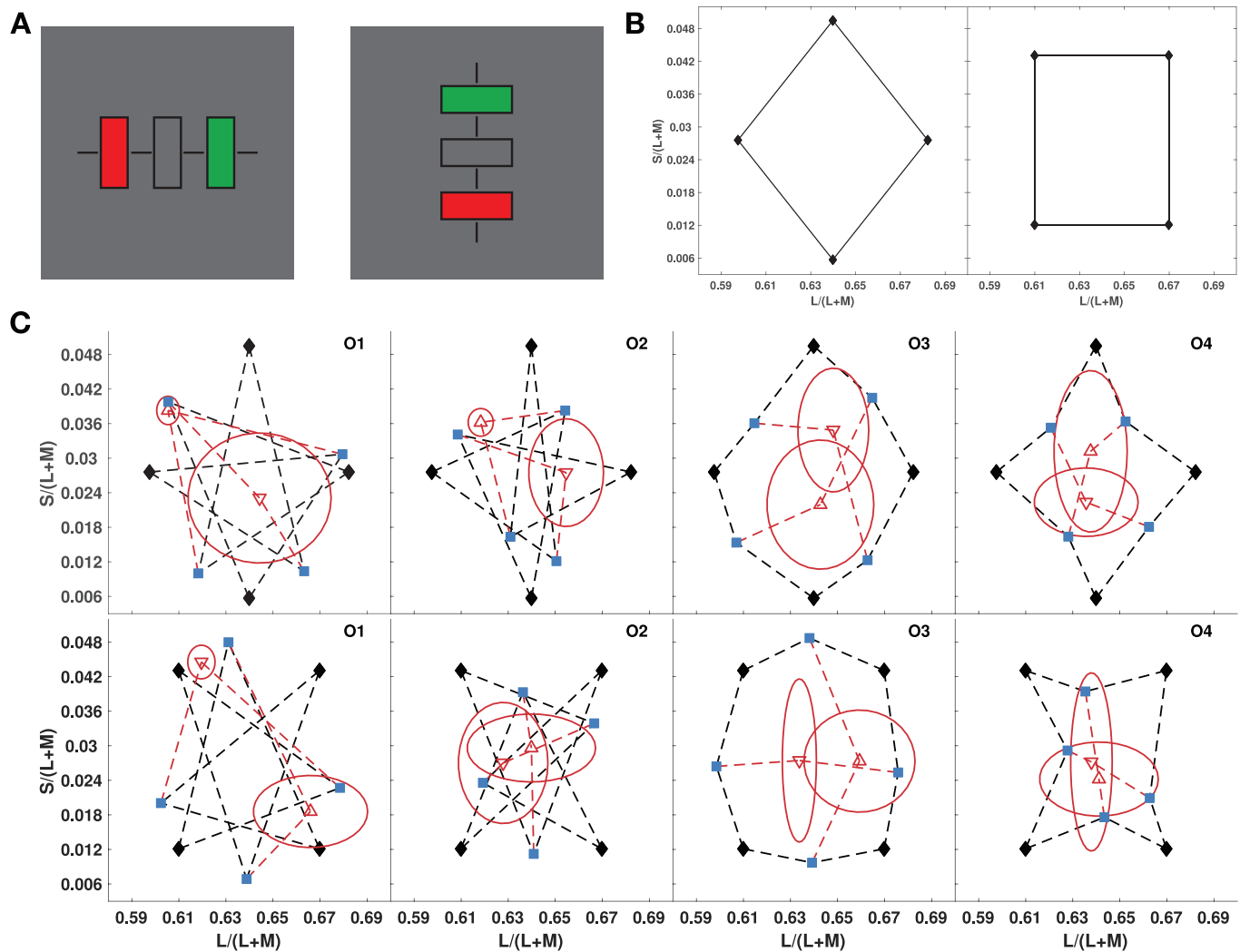


Figure 1. Tests for affine geometry of color similarity (Experiment 1). (A) Stimulus on each trial, alternated between vertical and horizontal on successive trials, to reduce the effect of aftereffects. Black bars assisted in central fixation. The colors of the flanking rectangles were fixed, and observers adjusted the color of the middle rectangle to their estimate of the midpoint color. (B) Color diamond and square used as the quadrilaterals (MacLeod-Boynton chromaticity diagram). (C) Results for four observers in Experiment 1 for the diamond (top row), and square (bottom row). Blue points show midpoint settings for the four sides of each quadrilateral, connected to their vertices by black dashed lines. Red points show midpoint settings for opposite blue points, connected by red dashed lines. Red ellipses depict standard deviations for the midpoint settings along the major and minor axes directions.

coinciding (Figure 1C), except for one case. Thus, perceptual color space is not Affine when observers set midpoints based purely on individual judgments of relative color similarity.

## Experiment 2: Effect of mental representations on intrinsic geometry

The results of Experiment 1 restrict color similarity space to at best a projective geometry, and this seems to

go against observers' abilities to reliably identify category boundaries between colors based on classifying colors as "reddish" versus "greenish," and "bluish" versus "yellowish" (Smithson & Zaidi, 2004). We thus repeated Experiment 1 with the same observers, but we gave them different instructions for finding the perceptual midpoint between the end-point colors: "Consider the colors of the flanking rectangles in terms of 'Red-Green' and 'Blue-Yellow' qualities. Next, judge the change in each of these qualities between the two colors. Then, set the central rectangle to the color that lies halfway in the 'Red-Green' interval defined by the flanking colors, and simultaneously half-way in the 'Blue-Yellow' interval." Observers were given no



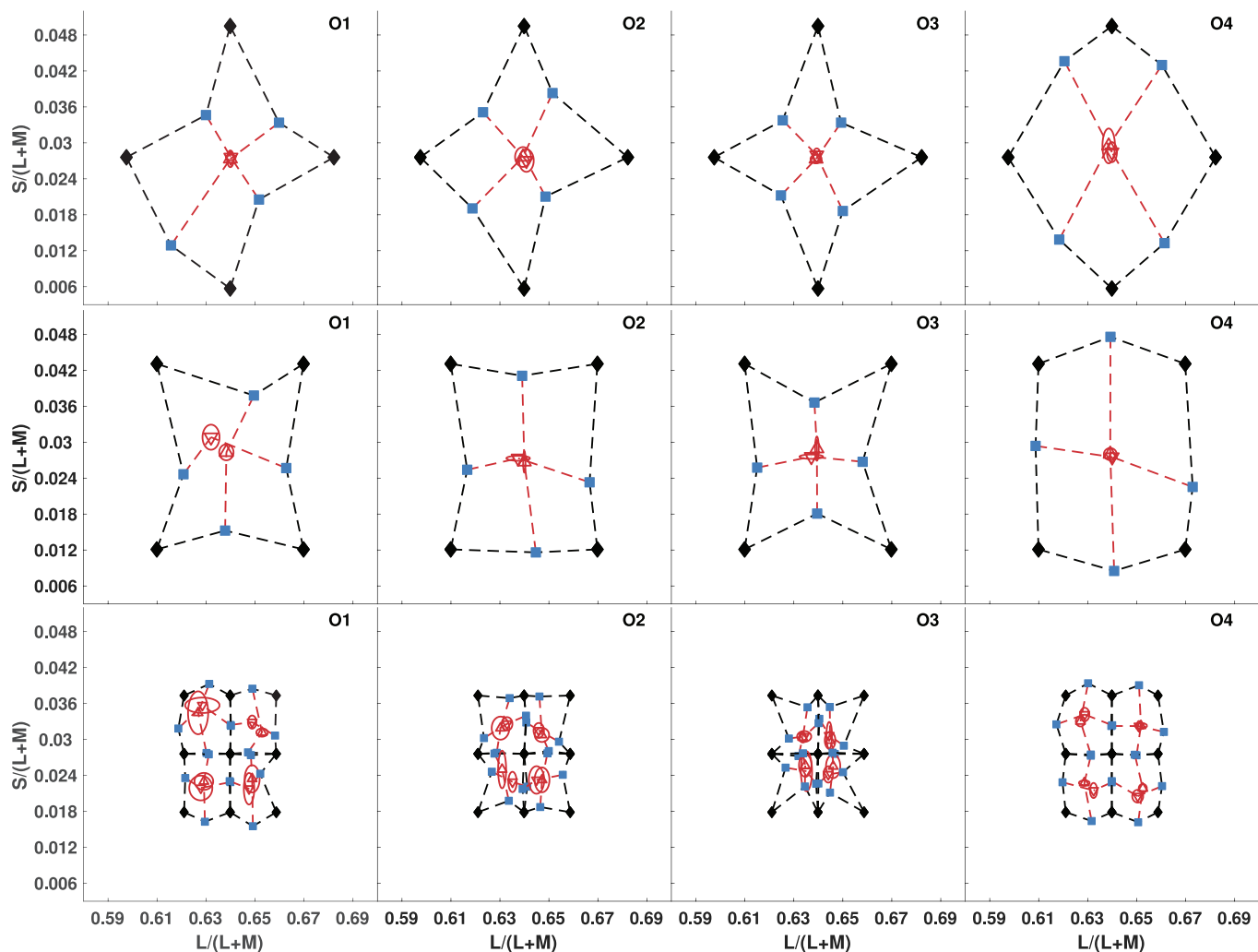


Figure 2. Testing for affine geometry of color similarity using opponent-color mental representations (Experiment 2): We tested the same two-color quadrilaterals as in Experiment 1 (top two rows), but with opponent-color instructions. We also tested four additional, smaller quadrilaterals, all of which shared a vertex at midgray (bottom row). Same plotting conventions as Figure 1C.

examples or definitions of “reddish”, “greenish”, “bluish”, or “yellowish”, or of the curvature of the opponent unique hue axes. The results of Experiment 2 in the left two columns of Figure 2 show that the midpoints estimating the centroid were coincident and lie near the center of each quadrilateral. Error ellipses in Experiment 2 for the same quadrilaterals as Experiment 1 are appreciably smaller. Consequently, similarity judgments made with opponent-color mental representations satisfy the conditions for an intrinsic affine geometry.

To further test the affine nature of the space, we also used four smaller squares, each of which had one vertex on midgray, instead of being centered on it (Figure 2, bottom row). The centroid estimating midpoints were again coincident within small errors, indicating that the intrinsic geometry is affine for midpoint judgments on both the large and small quadrilaterals.

### Experiment 3: Invariance of intrinsic color geometry to adaptation

If the geometric structure of color similarities revealed by Experiment 2 is an efficient representation of functionally important properties, we could expect it to be invariant to adaptation under different illuminations. We repeated Experiment 2, but with the stimuli illuminated by an additional light source. We used a Planar consisting of two LCD displays at  $110^\circ$  angle, superimposed via a 50/50 beam-splitter (Figure 3A). Two medium-sized color quadrilaterals were tested in this experiment, with one of nine adapting illuminants superimposed on the image of the test stimulus (Figure 3B). Figure 3B, right shows the coordinates of the test quadrilaterals with the added light from each of the illuminants. Superimposing the illumination from the

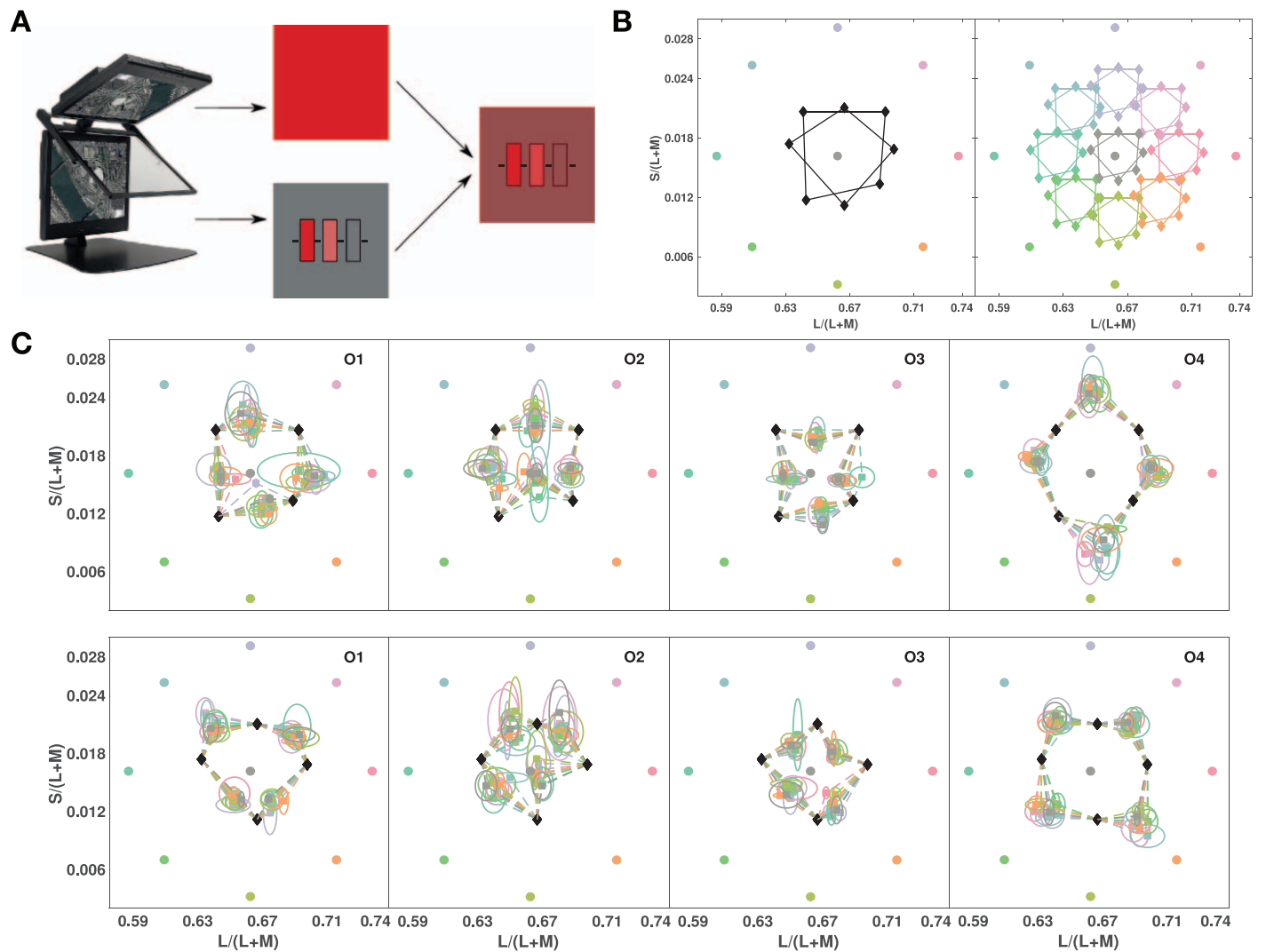


Figure 3. Testing for invariance of relative color similarities across adaptation conditions (Experiment 3): (A) Planar display with beam-splitter superimposing lights from two screens: A full-field red in the top display is combined with the test stimulus in the bottom, allowing test stimulus and task to stay unchanged under different adaptation conditions. (B, left) Chromaticities of the nine different superimposed full fields (in roughly corresponding colors) with the tested color quadrilaterals. (B, right) Chromaticity coordinates of the lights from test quadrilaterals superimposed with different full-field illuminants (in roughly corresponding colors). (C) Results plotted in the chromaticities of the test monitor for all full-field illuminants. Observers' chosen midpoints are shown color coded according to the illuminant under which they were measured and connected to their respective test vertices, top row for the square and bottom row for the diamond (black diamond symbols as in Figure 3B). Ellipses show the standard deviation of settings, color-coded for the full-field light.

top monitor on the stimulus from the bottom monitor is equivalent to adding the illuminant color vector to each test color. All the test vertices and the illuminants had the same luminance, so in Macleod-Boynton chromaticity space, this result shifts the quadrilaterals towards the illuminants. Because we were mainly interested in the transformations of perceptual color space that occur under different states of adaptation, observers set just the midpoints of the quadrilateral sides, adapting to one illuminant per experimental session.

Figure 3C shows the means of five repeated midpoint settings for each of nine different states of adaptation. The color of each midpoint symbol indicates the adapting illuminant used for that setting. For all observers, the different adapting illuminants had no appreciable influence on midpoint settings for either of the two tested quadrilaterals. Because the effect of a change in illumination spectrum is generally a vector addition, and the midpoint settings show little variation with adaptation state, their invariance can be explained by a subtractive adaptation process (see Discussion).

## Discussion

We demonstrate that the intrinsic geometrical structure of perceptual color space depends on the mental representation employed by an observer. Goodman (1972) pointed out that there are innumerable ways to assess similarity between two real or abstract entities, so the estimated degree of similarity depends entirely on the observer's perspective. Our results show that this observation also applies to color similarity. If observers judge similarity without being given an explicit representation, then colors seem to aggregate in categories without a natural ordering. This is not entirely unsurprising, because for somebody not schooled in color structure, there is no reason why white and green would not be judged as equally dissimilar colors from red. Color similarity has traditionally been studied by asking observers to rate the similarity of pairs of colors on a numerical scale, without any other instructions. The ratings are then subjected to multidimensional scaling, and the results presented in lower dimensional Euclidean spaces. It was therefore imperative to first have observers set midpoints by using their own notions of color similarity without being biased by instructions about a particular color representation; hence we have the design of Experiment 1. Note that observers were not forbidden to use any color representation they desired, so they could have on their own used notions of reddish versus greenish and yellowish versus bluish. That they did not use an opponent representation without being prompted to do so, indicates that that representation may not be used automatically in everyday judgments.

Instructions to use an opponent-color based representation revealed a perceptual color space with an Affine structure. This result allows us to compute ratios between segments of lines to estimate relative similarity between three colors on a line, and to predict observers' responses to parallel color changes, as would happen to colors of objects across different colored illuminants (Zaidi, 2001). The main effect of an opponent mental representation seems to be to locate white or gray in the center of the perceptual space, roughly at the intersection of the opponent dimensions. It is worth reiterating that “reddish”, “greenish”, “bluish”, and “yellowish” were judged by each observer only mentally and individually. The role of unique hues has been debated because there is no psychophysics or physiology showing primacy for the four unique hues, red, green, blue, and yellow (Wool et al., 2015). The results of this study suggest that they provide a possible systematic arrangement of colors that our observers were able to use almost effortlessly, much like a mental representation of the cardinal compass directions. Whether this arrangement would also be true for other pairs of complementary hues, or for speakers of languages that

do not have these color terms, remains to be seen. The results of such future studies would address the questions raised by Wittgenstein's (1977) assertion that concepts of red, blue, green, and other colors have meaning only because of their systematic interrelations, i.e., the language of color is a mathematical representation of a space akin to physical representation of particles in Euclidean coordinates.

This study was not designed to address the extrinsic geometry (curvature) of perceptual color space, but the midpoints also provide pairs of equal color similarities, so the chromaticities of the midpoints provide clues to the sign of curvature. A quadrilateral on a flat surface would bulge out if mapped on an ellipsoid, but on a hyperbolic surface, the sides of the quadrilateral would bend inwards and the internal angles would be less than 90°. With the opponent mental representation, the perceived midpoints of the quadrilateral sides lie consistently closer to the center for three observers, indicating an extrinsic hyperbolic geometry. One observer's midpoint settings, however, bulge outwards. Generally, each observer's midpoints for each side of the smaller quadrilateral follow a similar pattern as for the large quadrilaterals, suggesting the same class of extrinsic geometries for similarities across large versus small color separations. It seems that extrinsic geometries could be different across observers even if intrinsic geometry is the same. It may thus not be possible for a single nonlinear transform of a chromaticity space to represent similarities for all observers.

To test how well commonly used “uniform” color spaces, CIELUV and CIELAB, represent color similarities across the distances we tested, we plotted observers' midpoint settings in these spaces (Figure 4). In a truly uniform space, the perceived and calculated midpoints should coincide. For the large square and diamond quadrilaterals in Experiment 2, *t* tests ( $p < 0.05$ ) for four observers and eight midpoint settings revealed that 18 out of 32 measured midpoints deviated significantly from predicted midpoints in CIELAB, as did 18 out of 32 in CIELUV. Across the smaller quadrilaterals, the empirical midpoints were significantly different from predicted for 47 out of 128 in CIELAB and 33 out of 128 in CIELUV. Average departures of the midpoints for the larger quadrilaterals from the predictions were  $5.2 \pm 2.49$  (mean  $\pm$  *SD*) in  $\Delta E$  units in CIELAB according to CIEDE2000, and  $15.67 \pm 6.67$  (mean  $\pm$  *SD*) in  $\Delta E$  units in CIELUV. The empirical deviations from predicted midpoints may be marginally smaller than for MacLeod-Boynton space, but neither of these “uniform” spaces does an adequate job of representing color similarities between separated colors, and thus neither transform of CIE space provides an adequate representation of the extrinsic geometry of the color space. These spaces are probably adequate for delimiting industrial tolerance of

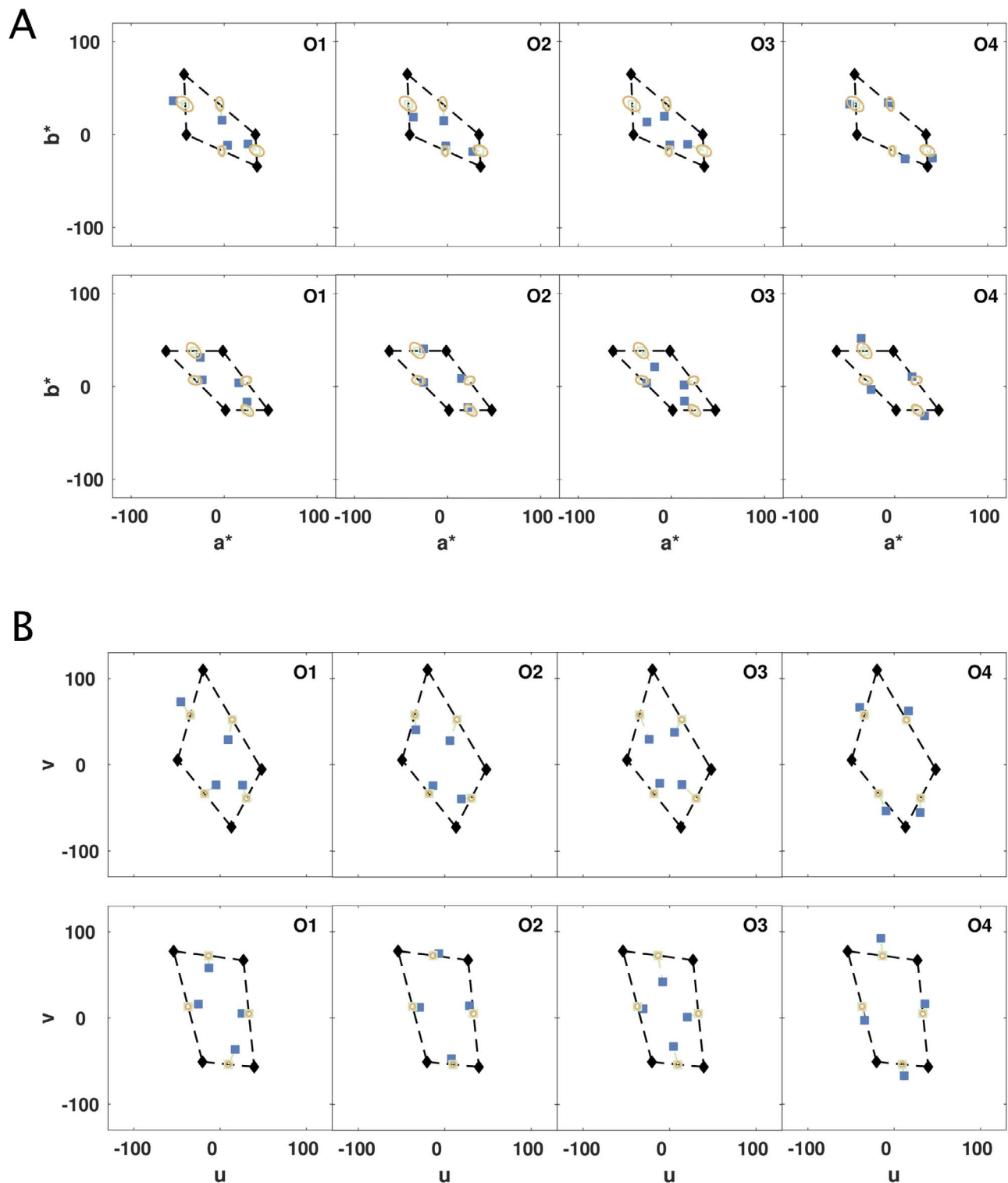


Figure 4. “Uniform” color spaces tested for perceived mid-points: CIELAB (A) and CIELUV (B). The data are from the top two rows of Figure 2, the diamond (top rows), and the square (bottom rows). The shapes are distorted by the nonlinear transformations of CIELAB and CIELUV. Blue squares show the average midpoint settings from each observer for the respective edge. The green empty squares show the predicted midpoint between the vertices. Thin green dashed lines connect the predicted midpoint with the blue setting of each observer. The yellow ellipses are 3 DE units, calculated according to CIEDE2000 in CIELAB and the standard Euclidean distance in CIELUV, and are centered on the predicted, green midpoints. Results in CIELAB ( $a^*$  vs.  $b^*$ ), and CIELUV ( $u$  vs.  $v$ ), and the statistical analyses in the main text show midpoint settings do not coincide with predicted midpoints.



color specifications but should not be used to estimate perceptual distances of separated colors, or to define a set of separated equally spaced colors.

Adapting to added full-field illuminants distributed about the hue circle did not affect midpoint settings, as if the visual system completely discounted the illuminant by subtracting it from the mixed stimulus. Superimposing the illuminant just creates an additive shift in the colors of the end points, and this shift retains the shape of each quadrilateral. Hence the discounting is reminiscent of Bergström's (1977) illumination and color analysis, which proposed that the visual system finds the common color vector across surfaces, and segments it from the scene to find the relative components of each reflecting surface. Although the process is compatible with a central process, the results could probably not rule out a quantitative model based on cone-adaptation either (Schnapf, Nunn, Meister, & Baylor, 1990). Based on electrophysiological results, however, it is most likely that the two minutes of adaptation to excursions from midgray to the "illuminant" color, prior to the measurements, were sufficient to evoke automatic subtractive adaptation mechanisms in ganglion cells of the retina (Zaidi, Ennis, Cao, & Lee, 2012), which counteract the additive shift. This early retinal adaptation transforms signals from the test stimuli so that they are not affected by the spectrum of the illumination; consequently, no adjustment is needed for later processes involving mental representations and hue judgments (Cicerone, Krantz, & Larimer, 1975) to obtain illumination invariant midpoints.

The adaptation invariance of geometrical color space provides a formal justification to generalize the color constancy measurements of Smithson and Zaidi (2004) to all of color space. They studied color constancy of patches simulating object colors in a variegated background under simulated sunlight and skylight. Completely adapted observers were asked to state whether the patch appeared "reddish" or "greenish", and "bluish" or "yellowish", thus providing estimates of color category boundaries. Plotted in terms of object reflectance, the category boundaries were found to be invariant to illumination change, indicating a high level of color constancy when observers are completely adapted to a single illuminant. However, invariance of object colors on category boundaries does not per se guarantee invariance of object colors within boundaries. The invariance of midpoint settings under different adaptation states in this study establishes that if locations of points on category boundaries are invariant, then since every interior point could be expressed as the midpoint of two boundary points, its location would also be invariant.

Observer midpoint settings were first used by Plateau (1872) to informally show that the gray midpoint

between white and black was essentially invariant to illumination and observer. This result restricts any hypothetical psychophysical scale that takes intensities to real numbers to be either a power or logarithmic function (Falmagne, 2002). A midpoint setting equates two similarities, and we show that color midpoint settings provide consistent and reliable estimates of relative similarity, and these are also invariant to illumination color. However, perceptual dimensions of color, such as hue, saturation, and brightness, are not independent, so deriving psychophysical scales for these dimensions requires considering their interactions.

At an abstract level, similarity is one of the fundamental principles used by Gestalt psychology to explain perceptual organization (Wertheimer, 1912) and can be used to generate models of generalization ranging from set-theoretic (Tversky, 1977) to continuous metric-space structures (Shepard, 1987), especially within a Bayesian formulation (Tennenbaum & Griffiths, 2001). Midpoint settings, representing equal relative similarities, can test the intrinsic geometry of a perceptual space as a logical preliminary to using multidimensional scaling, as they test the validity of the assumptions inherent in the statistical procedure. The methods in this paper are easily applied to other modalities, and could thus be used to critically test the geometrical structure of perceptual spaces that have been proposed on the basis of multidimensional scaling analyses for many other attributes, such as gloss, timbre, vowels, gestures, biological motion, tactile textures, tactile orientation, odors, and others (Zaidi et al., 2013).

Despite the extensive theoretical and empirical work on perceptual similarity, the neural basis of similarity computations is essentially unknown. Possibilities include activation patterns of receptors or later neurons, in rates or temporal patterns of impulse responses, and in different levels of correlated firing. The Affine geometrical structure we identify for color similarities suggests a simplification for neural circuits that compute similarity. Whereas calculating Euclidean distances in a perceptual space implies comparisons based on the power (sum of squares) of the difference, Affine geometry implies simpler comparisons based on ratios. It is possible that perceived colors are decoded using winner-take-all schemes on the responses of individual color-tuned IT neurons (Zaidi, Marshall, Thoen, & Conway, 2014). In that case, analyses of the population responses of IT cells may help us understand whether the restriction of perceptual color space to an affine geometry represents a trade-off between the goals of providing an efficient representation of sensory stimuli, and the costs of neural computations.

## Methods (Appendix contains more details)

### Observers

Four color-normal male observers, aged 27–32, gave written consent. This research was approved by the SUNY Optometry IRB.

### Data Analyses

Means and standard deviations of midpoint settings were the key statistics. Ellipses shown on the figures represent  $\pm 1$  *SD*.

### Experiments 1 and 2

Stimuli were presented on a calibrated HP1230 CRT (Hewlett-Packard, Inc., Palo Alto, CA), driven at 85 Hz by a Visage (CRS, Ltd.; Kent, UK) at 12 bits/gun. Colors in the MacLeod-Boynton (1979) chromaticity diagram were displayed using the procedure in Zaidi and Halevy (1993). Observers used an Xbox 360 controller (Microsoft, Corp., Redmond, WA) to set the target color. Three parallel  $4^\circ \times 0.6^\circ$  rectangles (Figure 1A) were placed at  $0.6^\circ$  separations on a midgray background. Each experiment was split into blocks of 20 trials each. There were 2 min of adaptation to the mean gray background before the first trial, and 2 s readaptation after each trial. The pairs of flanking colors on each trial were chosen by randomly sampling adjacent vertices from the tested quadrilaterals.

### Experiment 3

Stimuli were presented on a Planar SD2620W (Planar Systems, Inc., Hillsboro, OR) with two LCD displays at an angle of  $110^\circ$  superimposed by a beam-splitter (Figure 3A), each driven independently at 60 Hz by a dedicated port of an NVidia Geforce GTX 580 (NVidia, Corp., Santa Clara, CA). The displays had already been corrected for spatial distortions, color purity, and alignment (Jain & Zaidi, 2013) and were calibrated through the beam-splitter. The top monitor was used as a full-field spatially uniform illuminant, and its image was superimposed on the rectangles from the bottom monitor, simulating a lit scene. One of nine adapting illuminants was randomly chosen for each experimental session.

*Keywords:* perceptual space, color similarity, mental representation, perceptual geometry, color constancy, uniform color space

## Acknowledgments

There were NIH grants EY07556 and EY13312 to QZ. Thanks to Romain Bachy for extensive discussions about opponent-color representations, Jonathan Victor for discussions on perceptual spaces, and Bevil Conway for comments and suggestions on the paper.

RE and QZ designed the study, RE programmed and ran the experiments, RE and QZ analyzed the results and wrote the paper.

Commercial relationships: none.

Corresponding author: Qasim Zaidi.

Email: qz@sunyopt.edu.

Address: State University of New York, Graduate Center for Vision Research, NYC, NY, USA.

## References

- Arfib, D., Couturier, J. M., Kessous, L., & Verfaillie, V. (2002). Strategies of mapping between gesture data and synthesis model parameters using perceptual spaces. *Organised Sound*, 7, 127–144.
- Bensmaia, S. J., Denchev, P. V., Dammann, J. F., Craig, J. C., & Hsiao, S. S. (2008). The representation of stimulus orientation in the early stages of somatosensory processing. *The Journal of Neuroscience*, 28, 776–786.
- Bergström, S. S. (1977). Common and relative components of reflected light as information about the illumination, colour, and three-dimensional form of objects. *Scandinavian Journal of Psychology*, 18, 180–186.
- Brannan, D. A., Esplen, M. F., & Gray, J. J. (1999). *Geometry*. Cambridge, UK: Cambridge University Press.
- Chichilnisky, E. J., & Wandell, B. A. (1999). Trichromatic opponent color classification. *Vision Research*, 39, 3444–3458.
- Cicerone, C. M., Krantz, D. H., & Larimer, J. (1975). Opponent-process additivity-III. Effect of moderate chromatic adaptation. *Vision Research*, 15, 1125–1135.
- Cleland, T. A., Johnson, B. A., Leon, M., & Linster, C. (2007). Relational representation in the olfactory

- system. *Proceedings of the National Academy of Sciences, USA*, 104, 1953–1958.
- Coxeter, H. S. M., & Greitzer, S. L. (1967). *Quadrangles: Varignon's Theorem in geometry revisited* (pp. 51–56). Washington, DC: The Mathematical Association of America.
- Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in the lateral geniculate nucleus of macaque. *The Journal of Physiology*, 357, 241–265.
- Dimmick, F. L., & Hubbard, M. R. (1939). The spectral location of psychologically unique yellow, green, and blue. *The American Journal of Psychology*, 52(2), 242–254.
- Falmagne, J. C. (2002). *Elements of psychophysical theory*. Oxford, UK: Oxford University Press.
- Ferwerda, J. A., Pellacini, F., & Greenberg, D. (2001). Psychophysically based model of surface gloss perception. *Proceedings of SPIE 4299, Human Vision and Electronic Imaging VI*.
- Friele, L. F. C. (1961). Analysis of the Brown and Brown-MacAdam colour discrimination data. *Farbe*, 10, 193–224.
- Giese, M. A., & Lappe, M. (2002). Measurement of generalization fields for the recognition of biological motion. *Vision Research*, 42, 1847–1858.
- Goodman, N. (1972). *Seven strictures on similarity in problems and projects*. Indianapolis, IN: Bobbs-Merrill.
- Haanel, V. (2009). Pyoptical: Python interface to CRS OptiCAL. Retrieved from <https://github.com/esc/pyoptical>
- Hering, E. (1964). *Outlines of a theory of the light sense* (Hurvich, L. & Jameson, D. Trans.) Cambridge, MA: Harvard University Press. (Original work published 1920)
- Hollins, M., Bensmaia, S., Karlof, K., & Young, F. (2000). Individual differences in perceptual space for tactile textures: Evidence from multidimensional scaling. *Perception & Psychophysics*, 62, 1534–1544.
- Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review*, 64(6, Pt. 1), 384–404.
- Indow, T. (1980). Global color metric and color-appearance systems. *Color Research and Application*, 5, 5–12.
- Inglis, C. R., Jr. (1977). The spectral sensitivity of the opponent-color channels. *Vision Research*, 17(9), 1083–1089.
- Jain, A., & Zaidi, Q. (2013). Efficiency of extracting stereo-driven object motions. *Journal of Vision*, 13(1):18, 1–14, <https://doi.org/10.1167/13.1.18>. [PubMed] [Article]
- Kirchner, E. (2015). Color theory and color order in medieval Islam: A review. *Color Research and Application*, 40, 5–16.
- Klein, F. (1939). *Elementary mathematics from an advanced standpoint: Geometry*. New York, NY: Dover.
- König, A., & Dieterici, C. (1886). The modern development of Thomas Young's theory of colour-vision. *Report of the British Association for the Advancement of Science*, 56, 431–439.
- Krauskopf, J., & Gegenfurtner, K. (1992). Color discrimination and adaptation. *Vision Research*, 32, 2165–2175.
- Lakatos, S. (2000). A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics*, 62, 1426–1439.
- Larimer, J., Krantz, D. H., & Cicerone, C. M. (1975). Opponent process additivity—II. Yellow/blue equilibria and nonlinear models. *Vision Research*, 15(6), 723–731.
- Le Grand, Y. (1949). Les seuils différentiels de couleurs dans la théorie de Young [The differential color thresholds in Young's theory]. *Revue d'Optique Théorique et Instrumentale*, 28, 261–278.
- MacAdam, D. L. (1942). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America*, 32, 247–273.
- MacLeod, D. I. A., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, 69, 1183–1186.
- Marcu, G. (2003). Gray tracking correction for TFT-LCDs. *Proceedings of SPIE 5293, Color Imaging IX: Processing, Hardcopy, and Applications*.
- Maxwell, J. C. (1860). IV. On the theory of compound colours, and the relations of the colours of the spectrum. *Philosophical Transactions of the Royal Society of London*, 150, 57–84.
- McDermott, J. H., Schemitsch, M., & Simoncelli, E. P. (2013). Summary statistics in auditory perception. *Nature Neuroscience*, 16, 493–498.
- Pinto, N., Doukhan, D., DiCarlo, J. J., & Cox, D. D. (2009). A high-throughput screening approach to discovering good forms of biologically inspired visual representation. *PLoS Computational Biology*, 5: e1000579.
- Plateau, J. A. F. (1872). Sur la mesure des sensations physiques, et sur la loi qui lie l'intensité de ces sensations à l'intensité de la cause excitante [On the measure of physical sensation, and on the law that



- links the intensity of these sensations and the cause of excitation]. *Bulletins de l'Académie Royal des Sciences, des Lettres, et des Beaux-Arts de Belgique*, 33, 376–388.
- Pols, L. C. W., van der Kamp, L. J. T., & Plomp, R. (1969). Perceptual and physical space of vowel sounds. *The Journal of the Acoustical Society of America*, 46, 458–467.
- Radonjic, A., Cottaris, N. P., & Brainard, D. H. (2015). Color constancy supports cross-illumination color selection. *Journal of Vision*, 15(6):13, 1–13, <https://doi.org/10.1167/15.6.13>. [PubMed] [Article]
- Schnapf, J. L., Kraft, T. W., & Baylor, D. A. (1987, January 29). Spectral sensitivity of human cone photoreceptors. *Nature*, 325(6103), 439–441.
- Schnapf, J. L., Nunn, B. J., Meister, M., & Baylor, D. A. (1990). Visual transduction in cones of the monkey *Macaca fascicularis*. *The Journal of Physiology* 427(1), 681–713.
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. Part I. *Psychometrika*, 27, 125–140.
- Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820), 1317–1323.
- Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 nm and 500 nm. *Vision Research*, 15, 161–171.
- Smithson, H. E., Dinkova-Bruun, G., Gasper, G. E. M., Huxtable, M., McLeish, T. C. B., & Pantl, C. (2012). A three-dimensional color space from the 13th century. *Journal of the Optical Society of America A*, 29, A346–A352.
- Smithson, H., & Zaidi, Q. (2004). Color constancy in context: Roles for local adaptation and levels of reference. *Journal of Vision*, 4(9):3, 693–710, <https://doi.org/10.1167/4.9.3>. [PubMed] [Article]
- Stockman, A., & Sharpe, L. T. (2000). The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13), 1711–1737.
- Sun, H., Smithson, H. E., Zaidi, Q., & Lee, B. B. (2006). Specificity of cone inputs to macaque retinal ganglion cells. *Journal of Neurophysiology*, 95(2), 837–849.
- Tennenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24, 629–641.
- Terasawa, H., Slaney, M., & Berger, J. (2005). The thirteen colors of timbre. *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 323–326.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327–352.
- Victor, J. D., & Conte, M. M. (2012). Local image statistics: Maximum-entropy constructions and perceptual salience. *Journal of the Optical Society of America A*, 29, 1313–1345.
- Werner, J. S., & Wooten, B. R. (1979). Opponent chromatic mechanisms: Relation to photopigments and hue naming. *Journal of the Optical Society of America*, 69(3), 422–434.
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung [Experimental studies on the perception of movement]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, 61, 161–265.
- Wills, J., Agarwal, S., Kriegman, D., & Belongie, S. (2009). Toward a perceptual space for gloss. *ACM Transactions on Graphics*, 28, 103:1–103:15.
- Wittgenstein, L. (1977). *Remarks on colour* (L. L. McAlister & M. Schättle, Trans.). In G. E. M. Anscombe (Ed.). Oxford, UK: Blackwell.
- Wool, L. E., Kombar, S. J., Kremkow, J., Jansen, M., Li, X., Alonso, J. M., & Zaidi, Q. (2015). Salience of unique hues and implications for color theory. *Journal of Vision*, 15(2):10, 1–11, <https://doi.org/10.1167/15.2.10>. [PubMed] [Article]
- Wuerger, S. M., Maloney, L. T., & Krauskopf, J. (1995). Proximity judgments in color space: Tests of a Euclidean color geometry. *Vision Research*, 35, 827–835.
- Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae*. New York, NY: John Wiley and Sons.
- Zaidi, Q. (1998). Identification of illuminant and object colors: Heuristic-based algorithms. *Journal of the Optical Society of America A*, 15(7), 1767–1776.
- Zaidi, Q. (2001). Color constancy in a rough world. *Color Research and Application*, 26, S192–S200.
- Zaidi, Q., & Bostic, M. (2008). Color strategies for object identification. *Vision Research*, 48, 2673–2681.
- Zaidi, Q., Ennis, R., Cao, D., & Lee, B. (2012). Neural locus of color afterimages. *Current Biology*, 22, 220–224.
- Zaidi, Q., & Halevy, D. (1993). Visual mechanisms that signal the direction of color changes. *Vision Research*, 33, 1037–1051.
- Zaidi, Q., Marshall, J., Thoen, H., & Conway, B. R. (2014). Evolution of neural computations: Mantis



shrimp and human color decoding. *i-Perception*, 5, 492–496.

Zaidi, Q., Victor, J., McDermott, J., Geffen, M., Bensmaia, S., & Cleland, T. A. (2013). Perceptual spaces: Mathematical structures to neural mechanisms. *The Journal of Neuroscience*, 33, 17597–17602.

## Appendix: Varignon's Theorem and affine space

Varignon's Theorem states that for any quadrilateral, the figure formed by the midpoints of the four sides is a parallelogram. Since the diagonals of a parallelogram intersect at their midpoints, and the diagonals of the parallelogram are the bimedians of the quadrilateral, the point of intersection of the bimedians is simultaneously the midpoint of both lines (Figure A1). The chance of two random points coinciding is exceedingly small, hence the importance of this discovery. However, the theorem received another interpretation after the invention of vector algebra. If the vertices of the quadrilateral can be represented as vectors in an affine (or Euclidean) space, the two overlapping midpoints are represented by the same centroid vector, just calculated in different steps (Figure A2). The theorem can thus serve as a test of whether points in the space can be represented as vectors, subject to the rules of vector manipulation, with the noncoincidence of the midpoints refuting the conjecture that the space is at least affine.

## Appendix: Methods

### Observers

Four male observers, aged 27–32, participated in all three experiments. One of the observers was the author, RE. All observers had normal color vision in both eyes. If necessary, observers wore their corrective lenses. All participants were fully briefed on the purpose of the study at the completion of all experiments. The research was approved by and followed the policies of the SUNY Optometry IRB.

### Experiments 1 and 2

#### Equipment

Stimuli were presented on an HP1230 CRT (Hewlett-Packard, Inc., Palo Alto, CA). The HP1230 was

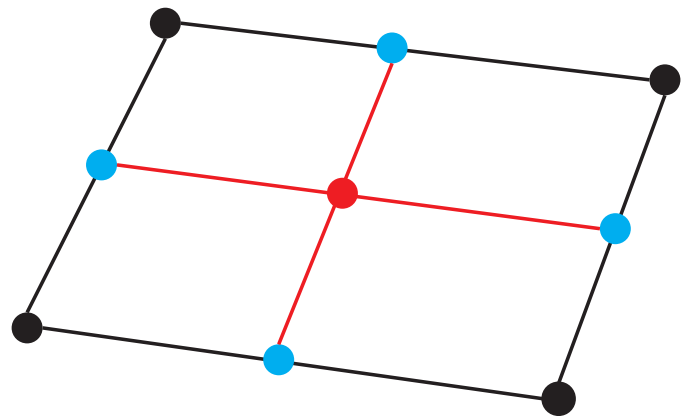


Figure A1. Varignon's Theorem illustrated for an arbitrary quadrilateral: The black lines and vertices denote the original quadrilateral. The blue points denote the midpoints on the four sides of the quadrilateral, and the red points (overlapping) denote the midpoints of the bimedians (red lines).

driven at 85 Hz by a Visage (CRS, Ltd.; Kent, UK) connected to a DELL Precision 380 (DELL, Inc., Round Rock, TX) running Windows XP Service Pack 3 (Microsoft, Corp.). It had a spatial resolution of  $1024 \times 768$ , minimum luminance of  $0.1 \text{ cd/m}^2$  and a maximum luminance of  $150.24 \text{ cd/m}^2$ . Stimuli for the HP1230 were programmed in MATLAB 2010a (MathWorks, Natick, MA) using the CRS Toolbox for MATLAB (v1.27).

Before calibration, we used the monitor's software menu to tune its settings for an optimal and accurate display of images (convergence, zero curvature of straight lines, color purity, etc.). The luminance curve for each phosphor of the HP1230 was measured using an OptiCAL (CRS, Ltd.; Kent, UK) via the VSG Desktop (v8.200) interface provided by CRS, Ltd. The monitor's voltage-versus-luminance curves, sampled at 256 levels for each gun, were fit with power curves using a least-squares procedure that was automated by the VSG Desktop. The chromaticity coordinates of the phosphors were measured with each at their maximum using a PR650 spectroradiometer (Photo Research, Inc., Syracuse, NY). The CIE1931 xyY coordinates of the phosphors were: red ( $x: 0.61, y: 0.340, Y: 34.69 \text{ cd/m}^2$ ), green ( $x: 0.280, y: 0.594, Y: 96.68 \text{ cd/m}^2$ ), and blue ( $x: 0.142, y: 0.070, Y: 20.94 \text{ cd/m}^2$ ). For all stimuli, colors were computed in the MacLeod-Boynton (MB) color space (MacLeod & Boynton, 1979; Derrington et al., 1984) and converted to RGB coordinates via a transformation matrix using the procedure reported in Zaidi and Halevy (1993). The CRS Toolbox's internal routines used the lookup table generated by the VSG Desktop calibration routine to properly transform the input and present the desired MB color on the HP1230.

Observers used an Xbox 360 (Microsoft, Corp.) controller to manipulate the target patch in each trial

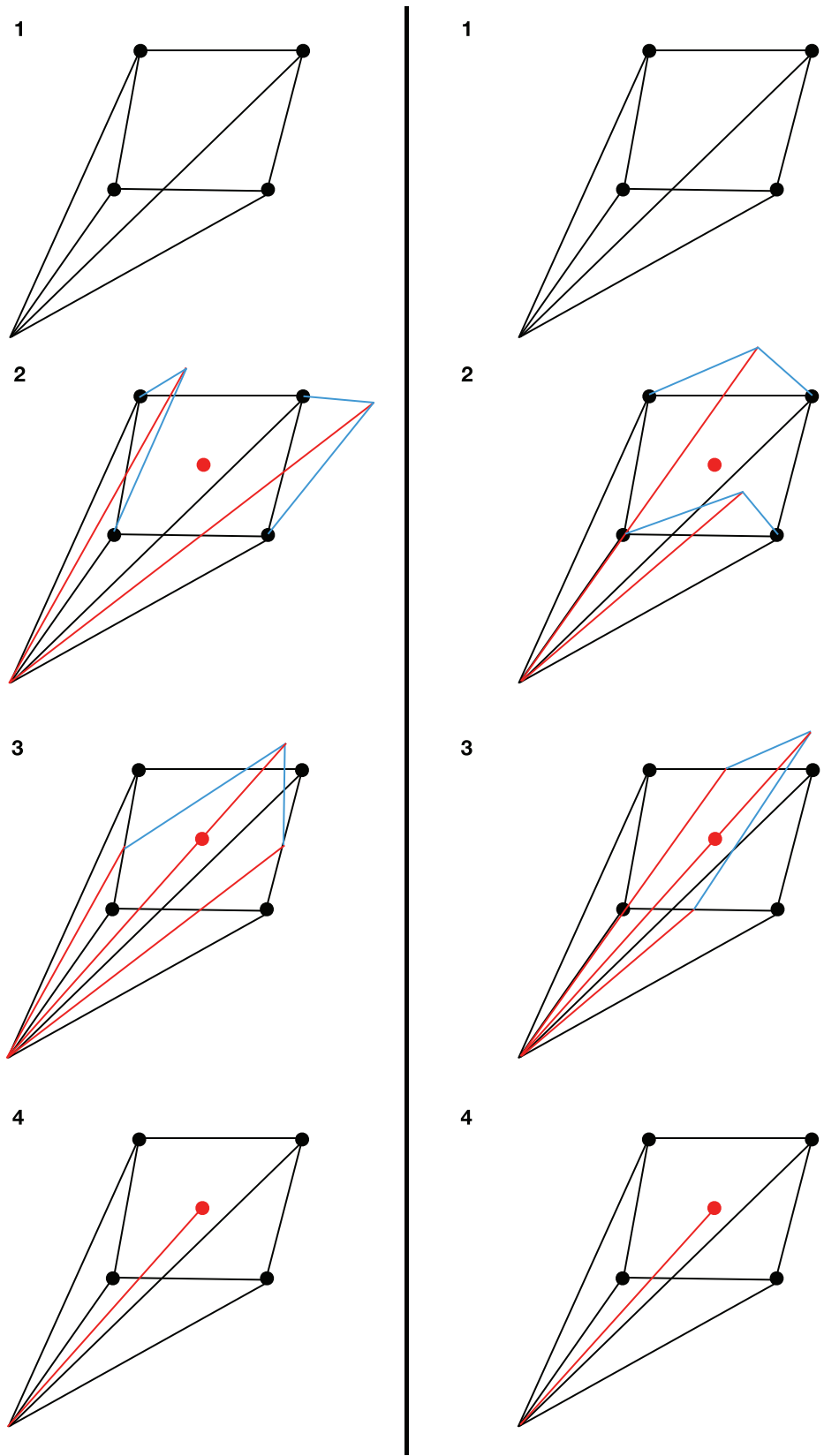


Figure A2. Varignon’s Theorem proven for vector spaces: (1) Vertices of the quadrilateral defined as vectors with arbitrary origin. (2) Two different opposite pairs of neighboring vectors are added. (3) The sum vectors are divided by 2 and the pairs of resulting vectors added. (4) The resultant vectors, divided by two, should be identical, as both give the centroid of the quadrilateral. Implication: If the vertices of a quadrilateral can be represented as vectors, the midpoints of both bimedians are the same centroid. Consequently, if the midpoints do not coincide, we can conclude that the vertices cannot be represented as vectors in an affine space, and thus not in a Euclidean space either.

and input their final response. The MATLAB joymex2 package was used to interface with the Xbox controller. The controller had two joysticks. Pushing the left joystick to the left or to the right changed the purity of the test patch and pushing the right joystick up or down changed its hue. These corresponded to changing the length and angular position, respectively, of a vector originating at midgray in the equiluminant plane of the MB space. When observers were satisfied with their midpoint setting, they pressed the “Y” button on the Xbox controller to continue to the next trial.

### Stimuli and task

All stimuli were viewed binocularly. The background was set to midgray. At the center of the monitor, three parallel  $4^\circ \times 0.6^\circ$  rectangles (Figure 1A) were placed side-by-side at  $0.6^\circ$  separations. The rectangles were either all oriented horizontally or all oriented vertically. The orientation switched on each trial and was not randomized in order to keep a balanced adaptation state across the visual field. The colors of the outermost flanking rectangles were fixed on each trial and were determined by randomly sampling neighboring vertices from color quadrilaterals specified in the equiluminant plane of the MB space (see Figures 1 and 2 in main text). The color of the central rectangle could be controlled by the observer. The central rectangle was always presented at the center of the screen. Each rectangle had a  $0.02^\circ$  (2 pixel) black border. Four black bars ( $0.02^\circ \times 0.6^\circ$ ) were presented orthogonal to the three rectangles and were vertically aligned along their centers to guide fixation. Otherwise, observers were allowed to freely fixate.

The goal of observers was to change the color of the central target rectangle until it appeared to be the midpoint color of the two flanking colored rectangles. They did this according to different instructions in Experiments 1 and 2. In Experiment 1, observers were instructed to use the joysticks to set the color of the central rectangle to the perceptual midpoint of the end-point colors by finding the color that was the most similar to both end-point colors, out of all colors equally similar to the end-point colors. In Experiment 2, observers were instructed to “Consider the colors of the flanking rectangles in terms of ‘Red-Green’ and ‘Blue-Yellow’ qualities. Next, judge the change in each of these qualities between the two colors. Then, set the central rectangle to the color that lies halfway in the ‘Red-Green’ interval defined by the flanking colors and simultaneously half-way in the ‘Blue-Yellow’ interval.” The adjustments in Experiment 2 could be considered generalizations of hue-cancellation measurements that are essentially memory matches to Hering’s unique hues (Hering, 1920/1964; Dimmick & Hubbard, 1939, Hurvich & Jameson, 1957; Larimer, Krantz, & Cicerone, 1975; Ingling, 1977; Werner & Wooten,

1979), Hue-cancellation requires titrating between percepts of reddish versus greenish (or yellowish versus bluish), but the titration is never to nonunique hues, as required for our task.

### Testing sequence

Each experiment was split into blocks of 20 trials each. The number of blocks per experiment depended on the number of color quadrilaterals, and the number of midpoint measurements. We measured 10 midpoint settings per edge for each quadrilateral and another ten midpoint settings for the bimedians. The first block was preceded by 2 min of adaptation to the mean gray background. After every trial, the relevant data were saved and a 2 s pause for readaptation to the background preceded the next trial. After every block, the observer was given the choice of either continuing the experiment or stopping. If the observer chose to continue, a 1 min readaptation period was initiated before the next block of trials began. If the observer chose to stop, the program saved the current state of the experiment. The next time the experiment was run, the program picked up from where the observer had stopped and initiated the next block of trials with 2 min of adaptation. The pairs of test colors on each trial were chosen by randomly sampling edges from the tested quadrilaterals. Once an edge was chosen, the two vertices that defined it were used as the colors for the flanking rectangles. The vertices of all quadrilaterals were identical for all observers for the first set of midpoint judgments. The second set of estimates of midpoints were made between pairs of midpoints estimated in the first set by each observer, so these vertices vary.

### Data analysis

Means and standard deviations of midpoint settings were used as the key statistics. Ellipses shown on the figures in the main text represent the standard deviation of settings along the axes of the MacLeod-Boynton diagram (i.e., the axes of the ellipses are aligned with axes of the diagram and their radius is the standard deviation of the settings in that direction).

## Experiment 3

### Equipment

Stimuli for Experiment 3 were presented on a Planar SD2620W 3D LCD Display (Planar Systems, Inc., Hillsboro, OR). The Planar had two LCD displays, each of which was independently driven at 60 Hz by a dedicated port of a NVidia Geforce GTX 580 (NVidia, Corp., Santa Clara, CA). The bottom of the two Planar monitors was situated perpendicular to the observer’s

line of sight. The top monitor was oriented  $110^\circ$  away from the bottom monitor, and the images that each monitor generated were superimposed via a beam splitter with a partly reflective coating that bisected the angle formed by the two monitors (fine adjustments were made to eliminate shadows at the edges of the display). Stimuli for the Planar display were programmed in MATLAB 2012a 32-bit (MathWorks) using Psychtoolbox (v3.0.9). The computer was custom built by iBUYPOWER (iBUYPOWER, Inc., Los Angeles, CA) and ran Windows XP Service Pack 2 64-bit edition (Microsoft, Corp.). The image from each of the Planar monitors was polarized for the presentation of 3D stimuli (the polarized light from the top monitor was rotated  $90^\circ$  relative to the polarized light of the bottom monitor), but we did not make use of this feature.

Observers viewed the Planar display without the 3D polarized glasses, since no 3D stimuli were displayed, and the glasses would have reduced the overall luminance reaching each eye. Rather, the top monitor of the Planar setup was used as a full-field illuminant and its image was superimposed on the image from the bottom monitor, simulating a lit scene. The view of the top monitor was blocked by a black piece of cardboard to prevent observers from viewing the illuminant and gaining any information about it. A black piece of cardboard was also attached to the bottom of the mirror, since it was possible to view part of the background on the bottom monitor without it. With both pieces of cardboard in place, the view of the stimulus was restricted to the superimposed illuminant-stimulus image.

The outputs of each Planar display were measured after each was filtered through the mirror. The Planar monitors were calibrated with the same PR650 that was used to calibrate the HP1230. Each of the Planar's two displays was calibrated while the other was turned off. A 1 s delay preceded each measurement with the PR650 to allow the LCD to settle into a stable state. We presented full screen patches on each display and calibrated both to have a similar midgray white point (CIE  $x$ : 0.329,  $y$ : 0.338) and minimum and maximum luminance (minimum =  $0.11 \text{ cd/m}^2$ , maximum =  $116 \text{ cd/m}^2$ ). We tested that the spectra from the primaries for each of the Planar displays overlapped and were indistinguishable by eye. The chromaticity coordinates of the three primaries were recorded at their maximum output and were checked for spatial/temporal stability and equality across the two displays. For the bottom display, the CIE1931  $xyY$  coordinates of the phosphors were: red ( $x$ : 0.6819,  $y$ : 0.3083,  $Y$ :  $35.34 \text{ cd/m}^2$ ), green ( $x$ : 0.2005,  $y$ : 0.6960,  $Y$ :  $79.03 \text{ cd/m}^2$ ), and blue ( $x$ : 0.1506,  $y$ : 0.0538,  $Y$ :  $7.28 \text{ cd/m}^2$ ). For the top display, the CIE1931  $xyY$  coordinates of the phosphors were: red ( $x$ : 0.6827,  $y$ : 0.3094,  $Y$ :  $32.77 \text{ cd/m}^2$ ), green ( $x$ : 0.2034,  $y$ : 0.6963,  $Y$ :  $79.38 \text{ cd/m}^2$ ), and blue ( $x$ : 0.1502,  $y$ : 0.0551,  $Y$ :  $7.71 \text{ cd/m}^2$ ). The chromaticity coordinates of the three guns on the LCD

were measured with the PR650 and used to construct a transformation matrix that took coordinates in the MB space to 8-bit RGB values, similar to the transformation matrix for the CRT setup. The Planar displays had already been corrected for spatial distortions, color purity, and alignment in a previous study (Jain & Zaidi, 2013). No deviations from these settings were found during calibration, so the monitors were left untouched. In particular, we ensured that the monitors were set to their factory standard display settings and were not emulating a showroom environment or a movie theater.

The input-output curves of each of the LCD primaries were measured for a square,  $6^\circ$  patch on a mean gray background that was stepped through the entire range of each primary, one at a time, while the other two primaries were set to their midlevel output. This was done to calibrate the LCDs to their outputs about the midgray, since LCDs are known to have poor gray tracking (Marcu, 2003), and all our equiluminant stimuli were essentially small-to-medium deviations from the midlevel output of each primary. Luminance measurements were taken with the same OptiCAL used to measure the HP1230, via a MATLAB-to-PyOptical interface (Haenel, 2009; free, open-source software developed by Valentin Haenel and modified for our lab setup by Martin Giesel). Lookup-tables (LUTs) were generated for each primary input-output curve (8-bit input = 0–255). The input (bit level) versus output (luminance) relationship for each primary, as measured with the OptiCAL, was fit with a spline via MATLAB's internal fitting routines. When a specific luminance for each gun was needed, the closest value in the LUT was found and the corresponding bit level was sent to the graphics card for display. The linearity of the gun outputs was verified, after the inputs had been calibrated via the LUT. The LCD input-output curves, after linearization, were fit with lines (all fits had  $R > 0.99$ ). The beam-splitter mirror was tested for linear additivity. We measured the output of each monitor individually at six different gray levels (0%, 10%, 25%, 50%, 75%, and 100%) through the mirror. Then, we measured the combination of both monitor outputs at each of these levels through the mirror. The superimposed gray levels showed no deviation from addition of the individually measured outputs. Observers used the two joysticks of the Xbox 360 controller to adjust the hue/saturation of the center rectangle, just as in Experiments 1 and 2.

### **Stimuli and task**

For this experiment, observers used the same instructions from Experiment 2 and saw similar stimuli, but now one of nine adapting illuminants was superimposed on the image of the three rectangles. The chromaticities of the quadrilaterals are slightly dis-



torted from the original diamond and square because of the different displays (Figure 3). Since we were solely interested in the transformations of perceptual color space that occur under different states of adaptation, only the midpoints of the sides were measured. The two quadrilaterals in Figure 3B of the main text were tested with these conditions. Since all observers had enough practice with midpoint settings, and midpoint settings were found to have small variability across trials in Experiment 2, only five measurements of each midpoint per edge of a color quadrilateral were taken, giving a total of 180 trials per observer.

### ***Testing sequence***

For Experiment 3, the adaptation states were randomly sampled, but only one adaptation state was tested during each experimental session. The test pairs for Experiment 3 were then randomly sampled in the same manner as in Experiment 2. The testing sequence was equivalent to Experiment 2, except that observers now adapted to the superimposed image of the mean gray background from the bottom monitor and the randomly chosen adapting illuminant from the top monitor.