

Aesthetic response to color combinations: preference, harmony, and similarity

Karen B. Schloss · Stephen E. Palmer

Published online: 10 November 2010

© The Author(s). This article is published with open access at Springerlink.com 2010

Abstract Previous studies of preference for and harmony of color combinations have produced confusing results. For example, some claim that harmony increases with hue similarity, whereas others claim that it decreases. We argue that such confusions are resolved by distinguishing among three types of judgments about color pairs: (1) preference for the pair as a whole, (2) harmony of the pair as a whole, and (3) preference for its figural color when viewed against its colored background. Empirical support for this distinction shows that pair preference and harmony both increase as hue similarity increases, but preference relies more strongly on component color preference and lightness contrast. Although pairs with highly contrastive hues are generally judged to be neither preferable nor harmonious, figural color preference ratings increase as hue contrast with the background increases. The present results thus refine and clarify some of the best-known and most contentious claims of color theorists.

Keywords Color preference · Color harmony · Color similarity · Color aesthetics · Color combinations

Introduction

Colors are rarely experienced in isolation. In nature, yellow daffodils are seen against green grass; in the built

Electronic supplementary material The online version of this article (doi:10.3758/s13414-010-0027-0) contains supplementary material, which is available to authorized users.

K. B. Schloss · S. E. Palmer
University of California,
Berkeley, CA 94720-1650, USA

K. B. Schloss (✉)
Psychology Department, University of California,
Berkeley, CA 94720-1650, USA
e-mail: kschloss@berkeley.edu

environment, a dark brown couch is viewed against a light beige wall; in Van Gough's *Starry Night*, the golden moon is highlighted against a deep blue sky. In all these examples, the aesthetic experience of any given color is strongly influenced by its participation in combinations of two or more colors. In discussing color aesthetics, it is therefore essential to consider not only how much people like individual colors (e.g., Hurlbert & Ling, 2007; Palmer & Schloss, 2010) but also how colors interact in more complex chromatic compositions.

In this article, we propose three distinct ways of evaluating perceptual responses to color combinations: (1) people's *aesthetic preference* for a given combination, (2) their perception of *harmony* for that combination, and (3) their preference for its *figural color* when viewed against a colored background. These concepts have often been confused and/or confounded in the literature on color combinations, as we explain below. We argue strongly for distinguishing among these three concepts, and show that they are demonstrably different when they are clearly defined and appropriately measured. Moreover, our results show that making these distinctions clarifies many previous confusions and resolves existing conflicts in the literature.

We define *pair preference* as how much an observer *likes* a given pair of colors as a Gestalt, or whole. We define *pair harmony* as how strongly an observer experiences the colors in the combination as *going or belonging together*, regardless of whether the observer likes the combination or not. These two judgments will be quite similar for an observer who likes harmonious color combinations (e.g., dark blue and light blue), but they can be arbitrarily different for an observer who likes contrastive color combinations (e.g., dark blue and saturated yellow). The distinction we draw between preference and harmony for colors is most easily understood by analogy to music. Nearly everyone who hears representative works by Mozart

and Stravinsky agrees that Mozart's music is more harmonious (or consonant) and Stravinsky's music is more disharmonious (or dissonant). Nevertheless, some people prefer Stravinsky, whereas others prefer Mozart. There will be a positive correlation between average judgments of musical harmony and musical preference if people generally prefer harmonious to disharmonious music, but that does not constitute evidence that they are conceptually the same. Because preference and harmony are so clearly different concepts in music perception, we are skeptical of claims that they are the same concept in color perception. Finally, we define *figural preference* as *how much the observer likes the figural color itself, when viewed against its background color*. Figural preference is only indirectly a measure of perception of the color combination because the observer is specifically asked to respond only to the figural color. It is nevertheless relevant to aesthetic response to color combinations because the same color can look quite different when viewed against different background colors, as documented in the well-known phenomenon of simultaneous color contrast (e.g., da Vinci, 1492; Chevreul, 1839; Helmholtz, 1866/1925; Walraven, 1976; Shevell, 1978).

Previous analyses of the aesthetics of color combinations have not clearly distinguished among these three types of judgments. "Preference" and "harmony" are often used interchangeably, and preference for a combination taken as a whole is frequently confused with preference for a figural color against a background color. For example, in one of the most influential art-based theories of color aesthetics, Chevreul (1839) used the terms 'preference' and 'harmony' as though they were equivalent, and further claimed that there are harmonies of both analogous colors and contrasting colors. Without going into detail, his *harmony of analogous colors* includes: (1) *harmony of scale* for colors that are similar in lightness and the same in hue and (2) *harmony of hues* for colors that are the same in lightness and similar in hue. *Harmony of contrast* includes: (1) *harmony of contrast of scale* for colors that differ significantly in lightness and are the same in hue, (2) *harmony of contrast of hues* for colors that differ in lightness and are similar in hue, and (3) *harmony of contrast of colors* for colors that are different in hue and different in lightness (although the lightness difference is claimed to be auxiliary). Other theories of harmony include Itten's (1973) theory that two or more colors are harmonious if they produce neutral gray when mixed together as paints, and Munsell's (1921) and Ostwald's (1932) theories that colors are harmonious when they have certain relations in color space (e.g., when they vary in lightness but are constant in hue and saturation), as well as other theories proposed by Nemcsics (1993), Goethe, 1810/2006, and Moon and Spencer (1944a, 1994b). (See Burchett, 2002 and Westland,

Laycock, Cheung, Henry, & Mahyar, 2007 for a review.) These theories are different enough that, if all their predictions were pooled, nearly every color pair could be considered harmonious!

The art theoretical literature is thus riddled with confusions and contradictions. Not surprisingly, these carry over to the empirical literature as well. For example, Granger (1952, 1953, 1995a, 1995b, 1995c) conducted an extensive series of experiments on color combinations but used "preference" and "harmony" interchangeably. Indeed, he inexplicably changes terminology from one article to another in the same issue of the same journal, referring to "harmony" judgments he reported in two of these articles (Granger, 1955a, 1955b), as "preferences" in the third (Granger, 1955c). Even so, it is useful to consider his tasks and results in light of the distinctions we raise among judgments of pair preference, pair harmony, and figural preference. Granger (1955a) found that perception of what he called "harmony" increased as hue difference increased. In Chevreul's terminology, this result appears to indicate that people perceive harmony of contrastive hues but not harmony of analogous hues. The task Granger (1955a) used, however, was ambiguous about what aspect of the color combinations was to be judged. He gave participants a color wheel with 20 removable hue wedges. Their task was to move one of the wedges (the "standard") around the circle until they found the hue "with which it made the best combination." When a hue was chosen, it was removed from the circle and the selection process was repeated until all the remaining hues were chosen, defining a rank ordering of the "harmonies" of each figural color against all background colors. In light of our three-fold distinction, it is manifestly unclear what criterion his observers should use to define the "best combination." Is it how well the colors go together (pair harmony), how preferable the combination is as a whole (pair preference), or which accompanying color made the standard color look best (figural preference)? Granger's (1955a) finding that "harmony" increased with increasing hue contrast resembles the pattern that we find when we ask observers to make ratings of figural preference (see Experiment 4) and the pattern Helson and Lansford (1970) found when they asked observers to rate "object colors" on different colored backgrounds. This suggests that Granger's (1955a) participants may actually have judged what we are calling figural preference: which accompanying (background) color made the standard (figural) color look best.

In the same journal issue, Granger (1955c) measured preferences and/or harmony again by asking participants to rank order single color preferences and all pair-wise combinations of 20 hues. He then modeled color combination preferences in terms of individual color preferences and hue distance. He found that harmony/preference increased as hue distance increased in this task as well,

suggesting that his subjects may actually have liked and/or found the *combinations* more harmonious when they differed greatly in hue. However, more recent empirical results (e.g., Ou & Luo, 2006, and those reported in [Experiment 1](#) below) have found the opposite. To make matters worse, Allen and Guilford (1936) measured the “affective value” of color combinations (presented side-by-side) and found no clear overall effect of hue similarity, although there was some evidence that very small or very large differences in hue were more pleasing than moderate differences. There has been additional empirical work on color harmony (e.g., Nemcsics, 2007, 2008, 2009a, 2009b), but it does not seem to settle the confusions described above.

A few previous art theorists (e.g., Albers, 1971) and perceptual researchers (e.g., Ou, Luo, Woodcock, & Wright, 2004a, 2004b) have made a distinction similar to the one we advocate between pair preference and pair harmony. Albers (1971), for example, argued against Chevreul’s idea that people necessarily prefer harmonious combinations, suggesting that dissonance can be as desirable as consonance. One can find evidence of this belief in many of his well-known color studies entitled “Homage to the Square.”

More recently, Ou et al. (2004a, 2004b) measured both preference and harmony for 190 color pairs by asking subjects to report two binary judgments: whether each pair was liked or disliked and whether it was harmonious or disharmonious. They found that average harmony and average preference judgments were indeed highly correlated ($r = +.85$), but emphasized that, even if an observer finds a pair to be harmonious, there is a moderate (31%) chance that he or she will dislike the color pair. However, Ou et al. (2004b) neglected to describe which types of combinations are harmonious yet disliked and to investigate whether there are individual differences in preference for harmony. In the present paper, we address both issues.

Thus far, we have focused on judgments of color combinations as a whole, either in terms of experiences of preference or harmony. Distinct from both of these judgments is preference for a figural color against a background color. Simultaneous color contrast is a well-known phenomenon: The color of the surround can strongly influence the appearance of the surrounded color (da Vinci, 1492; Chevreul, 1839; Helmholtz, 1866/1925; Walraven, 1976; Shevell, 1978). Presumably, this implies that the color of the background can also influence an observer’s preference for the figural color. Helson and Lansford (1970) studied the effects of background color on preference for “object” (figural) colors by asking participants to rate (from 1–9) 125 object colors against 25 different colored backgrounds. Object colors were more preferred against backgrounds with contrasting lightness and, to a lesser extent, contrasting saturation. The effects of hue difference were more

ambiguous but, generally speaking, object colors were more preferred on backgrounds with contrasting hues. It is noteworthy that, although Helson and Lansford (1970) framed their research question in terms of preference for “object colors” against different backgrounds—a clear example of *figural preference* in our terminology—they actually discussed their results in terms of pair preference and pair harmony without making a principled distinction among these types of judgments. Even so, it is clear from their description of the task that, in our terms, they were actually studying what we term figural preference for a foreground color against a colored background.

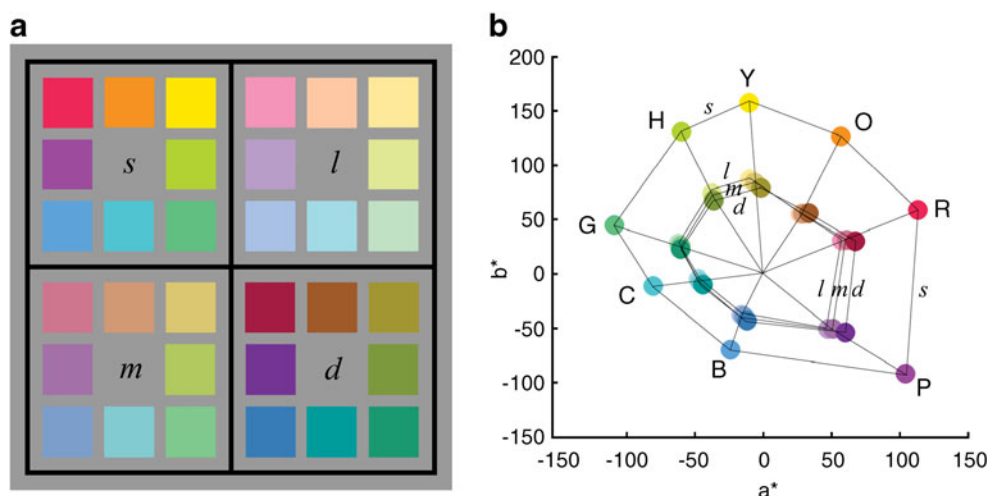
Camgöz, Yerner, and Güvenç (2002) also studied how background color influenced object color preference, but they reported finding no effects of similarity or contrast. This might have occurred because they only measured each participant’s single most preferred color on each of eight background hues, which is unlikely to have provided sufficiently detailed data to observe figural preference effects, even if they exist.

The experiments discussed here are part of the Berkeley Color Project (BCP), a massive repeated measures (MRM) design aimed at understanding color aesthetics within the context of color perception and various color associations (Palmer & Schloss, 2010). All participants completed the same set of 30 tasks (divided over 8 experimental sessions), using the same set of colors (see below) so that direct comparisons could be drawn across datasets. Palmer and Schloss (2010) previously reported BCP results on preference for single colors, showing that people like bluer colors better than yellower colors, have a higher preference for saturated colors relative to light and muted colors, and dislike dark yellow and orange more than any other colors (see “[General discussion](#) and conclusions” below for an explanation of these results). In the present paper, we will examine the same participants’ judgments of pair preference, pair harmony, and figural preference against colored backgrounds, drawing also on their single color preference ratings as assessed within the MRM design. We show that the three kinds of judgments distinguished above are empirically as well as conceptually distinct and that a principled analysis of their interrelations clarifies much of the confusion in the literature on perception of color combinations.

Experiment 1: Preference for color pairs

In [Experiment 1](#), we investigated preference for all pairwise combinations of the 32 chromatic colors studied in the BCP (see [Fig. 1](#)). The individual colors were the same as those reported in Palmer and Schloss (2010). They were sampled according to the dimensional structure of the Natural Color System (NCS) (Hård & Sivik, 1981), although they were actually chosen from the Munsell Book of Colors, Glossy Series (Munsell, 1966), and translated

Fig. 1 The 32 chromatic colors of the BCP as defined by eight hues, consisting of four approximately unique hues (*Red, Green, Yellow, Blue*) and their approximate angle bisectors (*Orange, chartreuse, Cyan, Purple*), at four “cuts” (saturation-lightness levels) in color-space (*Saturated, Light, Muted, and Dark*) (a) and the projections of these 32 colors onto an isoluminant plane in CIELAB color-space (Palmer & Schloss, 2010) (b)



into CIE xyY coordinates to generate them on our computer using the Munsell Renotation Table (Wyszecki & Stiles, 1967), as described in the Appendix. The sample included highly saturated colors of the four Hering primaries approximating the unique hues: red (R), green (G), blue (B), and yellow (Y), (Munsell hues 5R, 5Y, 3.75 G, and 10B, respectively). We also included four well-balanced binary hues that contained approximately equal amounts of the adjacent pair of unique hues: orange (O) between Y and R, purple (P) between R and B, cyan (C) between B and G, and chartreuse (H) between G and Y (Munsell hues 5YR, 5GY, 5BG, and 5P, respectively). We then defined four “cuts” through color space that differed in their saturation and lightness levels, as follows. Colors in the “saturated” (S) cut were defined as the most saturated color of each of the eight hues that could be produced on our monitor. Eight colors in the “muted” (M) cut were those that were approximately halfway between the S color and the Munsell value of 5 and chroma of 1 for the same hue. Eight colors in the “light” (L) cut were those that were approximately halfway between each S color and the Munsell value of 9 and chroma of 1 for the same hue. Eight colors in the “dark” (D) cut were those that were approximately halfway between each S cut and Munsell value of 1 and chroma of 1 for the same hue. The L, M, and D colors within each Munsell hue were equivalent in Munsell chroma (saturation). This set comprised the 32 chromatic colors that were studied (see Fig. 1).

Participants saw all possible pairs of the 32 chromatic colors in a figure-ground organization: a small square centered within a larger square, displayed against a neutral gray background. Both figure-ground organizations of each pair of colors were tested: A on B and B on A. For each pair, participants rated their aesthetic preference (how much they liked the pair as a whole) by selecting the appropriate point along a continuous line-mark response scale.

Methods

Participants There were 48 participants (24 males and 24 females) who completed all 30 tasks of the BCP. All participants were screened for color deficiency using the Dvorine Pseudo-Isochromatic Plates, and none of them were found to be color deficient. All participants gave informed consent, and the Committee for the Protection of Human Subjects at the University of California, Berkeley, approved the experimental protocol.

Design All pair-wise combinations of the 32 chromatic colors described above (see Fig. 1 and the Appendix) were used to generate 992 figure-ground color combinations.

Displays Test configurations were figure-ground pairs consisting of a small square (100×100 pix) centered on a larger square (300×300 pix). A continuous rating scale (400 pix long), containing demarcated center and endpoints, was located below the figure-ground pair. The rating scale was used to indicate how much each participant liked each display, ranging from “not at all” (written below the left endpoint) to “very much” (written below the right endpoint). Participants viewed the computer screen from approximately 70 cm. The monitor (Dell M990) was 18” diagonally with a resolution of $1,024 \times 768$ px. The background of the display was always a neutral gray (CIE $x = 0.312$, $y = 0.318$, $Y = 19.26$). The chromaticity and luminance functions of the red, green, and blue guns were measured as each gun ranged in voltage from 0–255 in equal steps of 17 using a Minolta CS100 Chroma Meter. The chromaticity and luminance functions for each gun were used to calculate the appropriate RGB values to ensure that we accurately presented the CIE xyY values for our colors. The displays were generated and presented using Presentation (www.neurobs.com).

Procedure The 992 figure-ground combinations were displayed one at a time in a random order. The participants' task was to indicate how much they *liked* each combination on a scale from “not at all” to “very much.” To respond, they used the mouse to move the cursor along the response scale and click on the point that best represented their degree of preference. Participants were informed that the vertical mark crossing the center of the scale represented a neutral (or zero) point. The recorded datum on a given trial corresponded to the x-coordinate (in pixels) at which the participant clicked on the scale for that trial, where 0 was the center of the scale. The response scale thus ranged from –200 (left endpoint of the 400-pix scale) to +200 (right endpoint of the 400-pix scale) and was normalized to range from –100 to +100 in the reported data. Trials were preceded by a 500-ms inter-trial interval (ITI) and lasted until participants made a response. Participants were allowed to take a break after each set of 60 trials.

Results and discussion

Mean preference ratings for color pairs as a function of figural hue and ground hue are plotted in Fig. 2a, averaged over S, L, M, and D cuts. The data show main effects of figural hue [$F(7, 329) = 8.32, p < .001$] and ground hue [$F(7, 329) = 10.70, p < .001$] as well as a powerful interaction between them [$F(49, 2303) = 25.42, p < .001$]. The same data, plotted in terms of hue angle in CIELAB color space, are available in the Supplementary Material (Fig. S10A). The pattern of results, although complex, is highly regular, with three primary features. First, the peaks in the functions of Fig. 2a show that figure-ground combinations for each ground hue are most preferred when the ground hue and figure hue are the same.¹ Second, pair preferences decrease monotonically as a function of the difference in hue between figure and ground. For example, the green ground hue function in Fig. 2a peaks when the figural color is another shade of the same green hue and decreases systematically as the figural color becomes less similar to green on both sides of the peak. (The reader is reminded that hue is a circular dimension, such that purple on the right end of the graph is perceptually similar to red on the left end of the graph.)

¹ Notice that there are no data from the conditions in which the figural hue is the same as the ground hue when both are in the same cut (lightness and saturation level), because there would be zero contrast between them. The statistical tests are therefore computed from the averages of all pairwise combinations of the four cuts for a given hue pair (16 pairs when the hues in the pairs are different and 12 pairs when the hues in the pairs are the same).

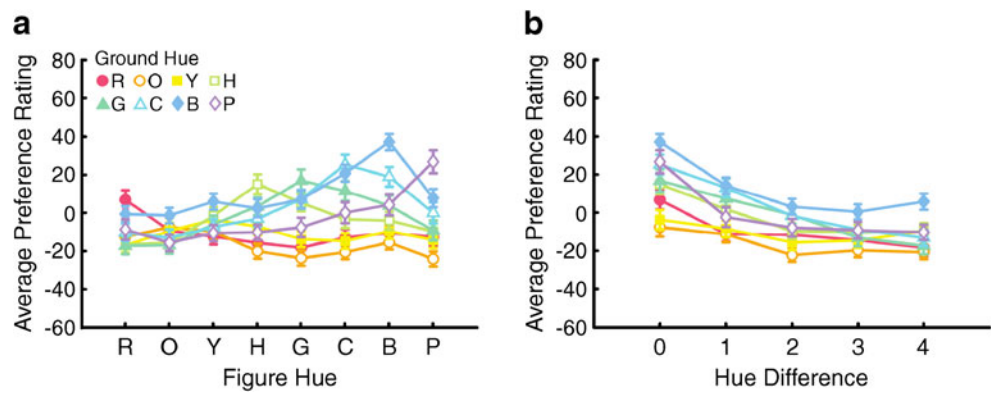
Figure 2b shows the same data as in Fig. 2a, but plotted as a function of the hue difference between the figure and ground colors (in terms of the number of hue steps in the BCP color sample). This plot emphasizes that pair preferences are highest when the figure and ground have the same hue (but differ in saturation and/or lightness levels) and decrease monotonically as hue difference between the figure and ground increases. It also provides clear evidence that people like Chevreul's (1839) “harmonies of analogous colors”, but virtually no evidence in favor of corresponding effects for contrastive hues. If the latter were present, the functions would curve upward toward the right end, where the figure and ground hues are maximally contrasting. No increases in preference for complementary colors are evident when the Bonferroni correction is applied to adjust for the eight *t* tests, one for each ground hue ($\alpha = .006$). The same data, plotted in terms of differences in hue angles in CIELAB color space, are available in the Supplementary Material (Fig. S10B).

Although this definition of “maximally contrasting” uses the perceptual complementary colors (red-green and yellow-blue), there is also little evidence of preference for contrastive hue combinations using paint-complementary colors: yellow-purple, blue-orange, and red-green. This was tested by comparing preference for pairs of paint complements versus the average of the pairs containing the two hues adjacent to their paint-complements [$F(1, 47) = 1.53, p > .05$], after accounting for the variance explained by figure and ground color preferences (when judged singly on a neutral gray background; see Palmer & Schloss, 2010). The only paint-complementary pair that was more preferred than its nearest neighbors (after applying the Bonferroni correction) was orange-blue compared with the average of orange-cyan and orange-purple [$F(1, 47) = 11.17, p < .008$].

The third salient feature of the results is the systematic variation in pair preferences with hue. Both the maxima of the ground-hue functions and their overall level generally increase as the hues become bluer and decrease as they become yellower. The strong correlation ($r = +.94$) between the level of the curves in Fig. 2b (mean preference across hue differences of 0–3) and the sharpness of their decline (slope of the best-fitting line between hue differences of 0–3) indicates that grounds containing more preferable hues (e.g., blue, cyan, and purple) get a larger preference increment when paired with figures of the same or similar hues than do grounds containing less preferable hues (e.g., yellow and orange).

Figure 3a isolates main effects of figural hue and ground hue. The shape of these functions, showing preference for cooler over warmer colors, closely resembles the shape of the hue preference function for single color preference ratings (Palmer & Schloss, 2010) from the same partic-

Fig. 2 Preference ratings for color pairs as a function of figural hue (*x-axis*) and ground hue (*separate lines*) (a) and as a function of the hue difference (in terms of steps in the BCP design) between the figure and ground (b). *Error bars* standard errors of the means (SEM)



ipants (Fig. 3b). This resemblance strongly suggests that preferences for color pairs are influenced to some degree by preferences for the component colors.

A multiple linear regression model was used to determine the degree to which the same participants' preferences for the component ground and figure colors (when judged singly on a neutral gray background; see Palmer & Schloss, 2010) could account for pair preferences. Only 21.7% of the variance in pair preferences for all 992 color pairs could be explained by single color preferences: 15% from ground color preference and an additional 6.7% from figural color preference. Ground color preference influences pair preference more than figural color preference, as indicated by the facts that ground color preference accounts for more variance than figural color preference and that the ground curve in Fig. 3a is more extreme than the figural curve. This somewhat surprising result may simply reflect the fact that the ground color covers more area than the figural color. Even so, this additive model based on single color preferences accounts for relatively little variance in the overall pattern of results because it cannot, by definition, explain the complex figural-hue \times ground-hue interaction so clearly present in Fig. 2a. One or more relational factors are required. Below, we attempt to identify what those relational factors might be using various predictors derived from Munsell dimensions.

The ten Munsell factors considered in this analysis were the hue difference (the number of Munsell hue steps by which the figural and ground colors differed), the sum, the signed difference, and the absolute value of the figure-ground difference in hue coolness (the number of Munsell hue-steps removed from Munsell hue 10R)², the value (or lightness) and the chroma (or saturation) of the figural and ground colors. All possible combinations of factors were tested for

all possible numbers of factors (i.e., all pairs of factors were tested in 2-factor models and all triplets were tested in 3-factor models, and so on up to 10 factors). The model we report as the "best" model was the model that explained the largest percentage of variance and that also explained at least 1% more variance than the next best model with the same number of factors. We also report the results of the "full model" that includes all factors, but we do not name or give the order of entry for the factors included beyond those in the best model as just defined.

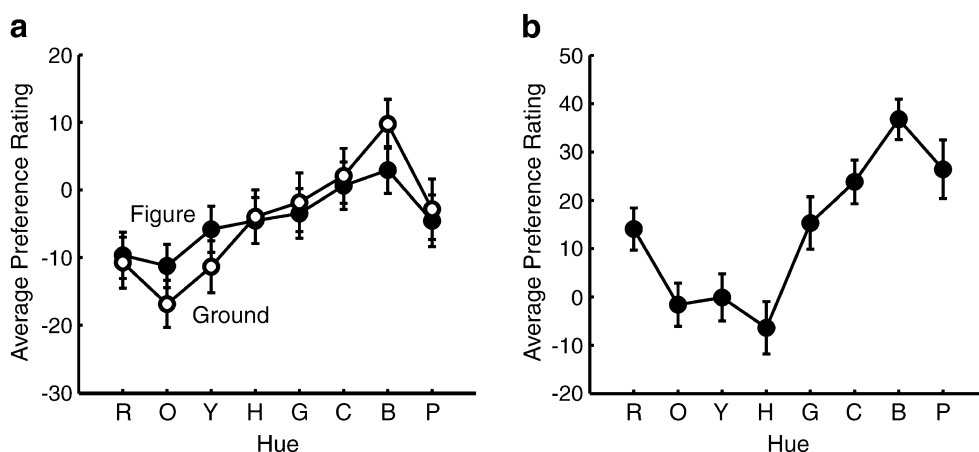
The left-most bar in Fig. 4 shows the best fitting model for pair preference ratings, where each factor's increment in percentage of variance explained is represented by a corresponding increment in the height of the bar, with the lowest segment being the factor that was entered first. The best fitting model explained 53.5% of the variance in pair preference ratings, showing that more preferred pairs contained cooler colors that were similar in hue and contrasting in Munsell value (lightness). An additional 7% of the variance can be explained in the full model when all 10 factors are included, but there is no clearly defined "best" model (see above) in any of the regressions containing more than 3 factors.

When figure and ground color preference are added to the 3-factor Munsell regression model shown in Fig. 4, they account for an additional 9.4% of the variance (6.9% from ground color preference and 2.5% from figural color preference). This brings the total amount of variance explained to 62.9%, which shows that component color preferences are still important after the variance due to the relational factors in the Munsell model has been removed. In discussing the results of Experiment 2, however, we report an even better model, which explains 80.8% of the variance, based on rated color harmony as a relational factor.

To further understand the nature of pair preferences, we also examined the effects of figural and ground cut: saturated (S), light (L), muted (M) and dark (D). The means that were analyzed (see Fig. 5) only included pairs with hue-difference steps of 1 through 4 because there were

² 10R (red-orange) was chosen because it was closest to the minimum of the coolness function obtained from participants' ratings of this dimension.

Fig. 3 Main effects of ground hue (*open circles*) and figure hue (*closed circles*) for pair preference ratings (**a**), and for single color preferences of the same participants (Palmer & Schloss, 2010) (**b**). Error bars standard errors of the means (SEM)



no zero hue-difference data for same-cut pairs. The results show no main effects of figural cut [$F(3, 141) = 2.90, p > .05$] or ground cut ($F < 1$), but there was a reliable interaction between them [$F(9, 423) = 7.66, p < .001$]. Pair-wise comparisons of cut combinations showed that the only effects of cut occurred for the saturated ground conditions: Combinations with saturated figures on saturated grounds were preferred to those with light, muted and dark figures on saturated grounds [$t(47) = 3.74, 6.33, 3.56, p < .002$], and those with light figures on saturated grounds were preferred to those with muted figures on saturated grounds [$t(47) = 3.64, p < .002$]. (A critical value of .002 was used

after applying the Bonferroni correction to compensate for the 24 comparisons.) The effects of figure and ground cut as a function of hue difference between the figure and ground colors can be found in the Supplementary Material (Fig. S1). These data indicate that pair preferences decrease as the hue difference between figure and ground colors increase for all cut-combinations, consistent with the inclusion of contrast in Munsell lightness (value) in the previously described best-fitting regression model.

We also examined figure-ground asymmetries in preference (e.g., warmer-figure/cooler-ground vs. cooler-figure/warmer-ground) to see whether figure-ground status influ-

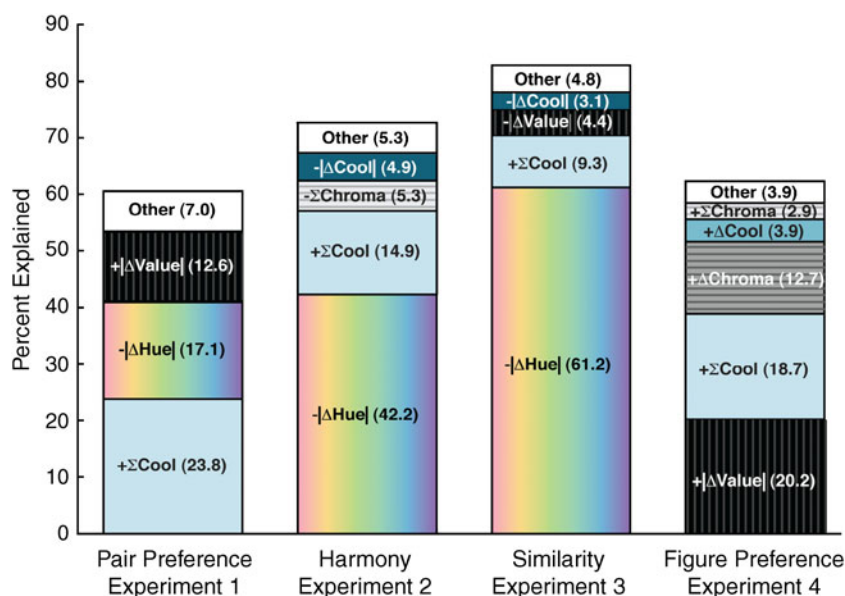


Fig. 4 Bars show the percentages of variance explained by the best-fitting Munsell models for pair preference (Experiment 1), pair harmony (Experiment 2), two-color similarity (Experiment 3), and figural color preference against colored backgrounds (Experiment 4). Stripes within each bar show the percentage of variance explained by each factor in the order with which they were entered in to the regression model (bottom to top). The sign before each term indicates

whether the factor was positively or negatively weighted in the corresponding regression equation (e.g., “+ΣCool” indicates that the sum of the coolnesses of the component colors was positively related to rated preference, harmony, and similarity, and “-|ΔHue|” indicates that the absolute value of their difference in hue was negatively related to these ratings)

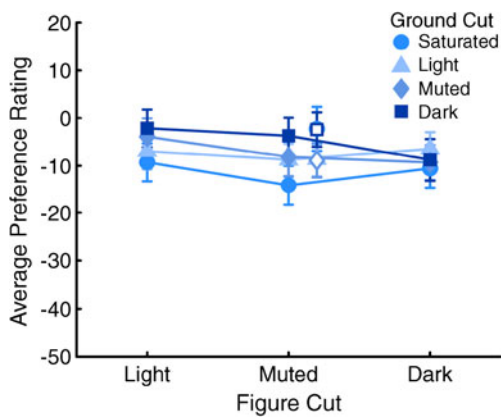


Fig. 5 Preference for color pairs for each ground cut (*separate lines*), as function of figure cut (*x-axis*). Data points for the saturated (*S*) figure cuts (*open symbols*) are plotted separately at an *x-axis* level similar to the muted (*M*) colors because they share similar lightness levels, but are slightly offset for clarity. *Error bars* standard errors of the means (SEM)

enced pair preferences by testing the signed difference between the figure and ground color along the Munsell dimensions tested above. Pair preferences were slightly, but significantly, correlated with the differences in coolness, such that pairs with warmer figures on cooler grounds were preferred to the reverse ($r = +.13$, $p < .001$). The same was true of differences in Munsell value: pairs with lighter figures on darker grounds were preferred to the reverse ($r = +.14$, $p < .001$). Nevertheless, these differences due to spatial figure-ground organization were quite small in comparison with the differences due to different colors. A regression model based on these two spatial predictors explained only 4% of the variance in pair preference, with the value difference accounting for 2% of the variance (lighter figures being preferred) and coolness differences accounting for an additional 2% (warmer figures being preferred). A further investigation of preference asymmetries using a two alternative forced choice task, in which the only difference between the two pairs in the comparison was the figure-ground assignment of the colors, will be presented in a subsequent paper (Schloss & Palmer, 2010). Preliminary results show that the asymmetries of coolness and lightness noted here are robust in the forced choice task.

Experiment 2: Color harmony and its relation to preference

In [Experiment 1](#), we showed that there are clear, systematic patterns in preferences for color pairs that are governed primarily by component color preferences, coolness, hue similarity, and lightness contrast. In [Experiment 2](#), we investigate what factors influence color harmony ratings and how they relate to pair preference ratings.

Findings previously reported by Ou and colleagues (Chuang & Ou, 2001; Ou et al., 2004a, 2004b; Ou & Luo, 2006) suggest that perceived harmony of color pairs is closely related to pair preference. Chuang and Ou (2001) found that pairs in which both colors were the same in hue were judged as more harmonious than those with different hues, and we found the same to be true for pair preferences in [Experiment 1](#). They also found that pairs that were different in luminance were judged to be more harmonious than those that were similar in luminance, and we found the same to be true for pair preferences in [Experiment 1](#). They further reported that preference for the component colors of a pair influenced harmony judgments: pairs that included two favorite colors were most harmonious, followed by pairs that included one favorite color and then pairs with no favorite color. Ou and Luo (2006) later reported that pairs were harmonious when colors were similar in hue, different in lightness, had a high combined (summed) lightness, and included light yellow as a component. Unfortunately, many of these conclusions are compromised by Chuang and Ou's definitions of harmony as "that which pleases the viewer" or "that which is harmonious." In the first definition, it is unclear whether "pleasing" refers to how well the colors go together (what we call pair harmony) or how much the observer likes the pair (what we call pair preference). Their second definition of harmony is simply circular and thus meaningless.

Our primary goal for [Experiment 2](#) was to obtain harmony ratings that were uncontaminated by confusions with pair preference using the same participants and the same colors as in [Experiment 1](#). We then used these ratings to determine how well people's harmony judgments can explain the pattern of variation in their pair preferences (see [Experiment 1](#)). In particular, we suspected that perceived harmony might be the relational variable that would best complement preferences for the component figure and ground colors in explaining people's preference ratings for color pairs. We also sought to reexamine the findings of Chuang and Ou (2001) using a more refined definition of harmony by including the musical analogy described in the introduction when instructing our observers about the difference between harmony and preference. Secondly, we wanted to examine individual differences in "preference-for-harmony" as indexed by the correlation between people's pair preference ratings in [Experiment 1](#) and their harmony ratings in [Experiment 2](#).

Methods

Participants The participants were the same 48 observers who completed [Experiment 1](#).

Design and Displays The design and displays were the same as in Experiment 1, except that the left endpoint of the rating-scale line was labeled “dissonant” and the right endpoint was labeled “harmonious.”

Procedure As in Experiment 1, participants were presented with each of 992 chromatic figure-ground combinations, one at a time in a random order. The harmony task was to indicate how “harmonious” the figure-ground color pair was on a scale from “dissonant” to “harmonious.” In order to clarify the difference between preference and harmony, participants were told the following: “Your task will be to indicate how “harmonious” you find each combination—how well the colors go together—by clicking a point on a scale like the one below. We are not asking you to rate how much you like each pair of colors. Some people like color combinations that are harmonious and others like combinations that are dissonant. For example, in music, some like Mozart and others like Stravinsky, but everyone would agree that Mozart is more harmonious and Stravinsky is more dissonant.” The harmony-rating task was completed in a different testing session that took place after the preference-rating task.

Results and discussion

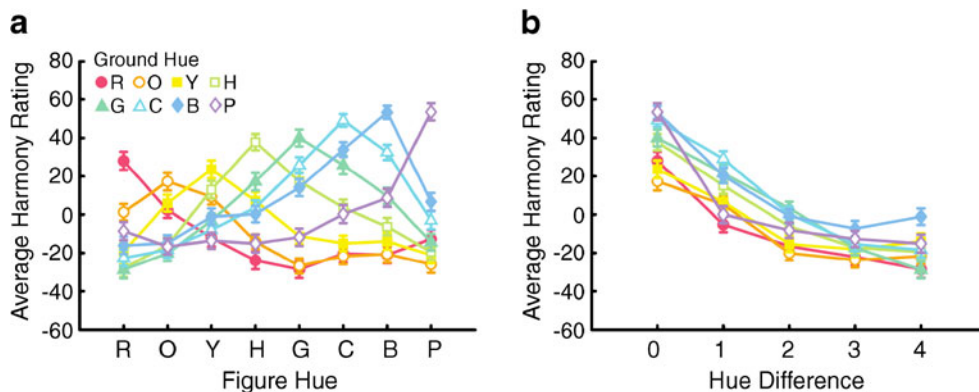
Because Chuang and Ou (2001) reported that their harmony data were influenced by preferences for the component figure and ground colors, we specifically tailored our instructions to try to dissociate such effects. To examine the extent to which we succeeded, we first examined the influence of figure preference and ground preference on harmony ratings in a two-factor regression analysis. The results show that only 1.4% of the variance in our harmony ratings is due to component color preferences: 1.1% from ground color preference and an additional 0.3% from figural color preference. This amount is an order of magnitude less than the 21.7% of the variance that is due

to figure preference and ground preference in the pair preference data of Experiment 1. This striking reduction supports our contention that, with appropriate instructions, observers can make harmony ratings that are essentially unaffected by their single color preferences. This difference between the present results and those of Chuang and Ou (2001) also supports our belief that their observers probably interpreted their instruction to judge how “pleasing” the color pairs were as asking, to some extent, about preference rather than or in addition to harmony (at least as we defined it in our instructions).

The pattern of color harmony ratings as a function of figural hue and ground hue is shown in Fig. 6a. Notice first that it is strikingly similar to the pattern of results for pair preference ratings but somewhat more exaggerated. Indeed, the correlation between average pair-wise preference ratings and average pair-wise harmony ratings was +0.79, accounting for 62% of the variance. Given this strong positive relation, it is understandable that Chevreul and other color theorists erroneously equated color harmony and color preference: generally speaking, people do tend to prefer harmonious color combinations. That does not mean that harmony and preference are either conceptually or empirically the same, however. It is also noteworthy that there was greater agreement among participants about their judgments of pair harmony than about their judgments of pair preference. The correlation of each observer’s harmony ratings with the group-average harmony ratings (average $r = +.51$) was significantly greater than the corresponding correlation of their preference ratings with the group-average preference ratings (average $r = +.36$) [$t(47) = 5.72, p < .001$]. This fact indicates that, whatever perceived color harmony might be, people are in better agreement about it than about their preferences for the same colored displays. The same data, plotted in terms of hue angle in CIELAB color space, can be found the Supplementary Material (Fig. S11A).

The harmony data in Fig. 6a reveal main effects of both figural hue [$F(7, 329) = 28.92, p < .001$] and ground hue [$F(7, 329) = 22.80, p < .001$], as well as a strong interaction

Fig. 6 Harmony ratings for color pairs as a function of figural hue (x-axis) and ground hue (separate lines) (a) and as a function of the hue difference (in terms of steps in the present BCP design) between the figure and ground (b). Error bars standard errors of the means (SEM)



between them [$F(49, 2303) = 64.85, p < .001$]. Harmony ratings were highest for each pair when the figure and ground hues were the same, and they decreased monotonically as hue difference increased. This result is consistent with Chevreul's (1839) claim that analogous colors are harmonious. It is also consistent with previous empirical studies of color harmony in which harmony was defined as "pleasantness" (e.g., Chuang & Ou, 2001; Ou & Luo, 2006), even though the latter data appear to be contaminated by single color preferences for the reasons outlined above.

As was also true for pair preferences, there is virtually no evidence supporting Chevreul's (1839) claim that contrastive hues are harmonious. If there had been, the harmony curves in Fig. 6b, which are plotted as a function of hue difference, would curve upward toward the right end, where the figure and ground hues are maximally contrasting (red-green and blue-yellow). Instead, when these data are averaged over ground hue, there is a reliable decrease in harmony ratings for pairs from the hue-step 3 to hue-step 4 conditions [$F(1, 47) = 6.11, p < .05$]. The same is true for hues paired with their paint-complement (blue-orange and yellow-purple): paint-complement pairs were rated as reliably less harmonious than the same hues paired with the average of the two hues adjacent to their paint-complement [$F(1, 47) = 17.67, p < .001$]. Thus, the results are not in accord with what Chevreul presumably would have predicted. The only reliable up-turn is for the blue-ground/yellow-figure combination [$F(1, 47) = 11.05, p < .006$], which may be an artifact arising from the fact that blue and gold (essentially, a shade of yellow) are the official school colors of the University of California, Berkeley, where the experiments were conducted. (See Schloss, Poggesi, & Palmer, 2010, for an in-depth study of the influence of school colors on the color preferences of Berkeley and Stanford students.) The reliable increment for blue and yellow combinations over their immediately adjacent neighbors may also be due to the large lightness contrast between them. The same data, plotted in terms of differences in hue angles in CIELAB color space are available in the Supplementary Material (Fig. S11B).

An analysis of the effects of cut (saturation/lightness level) showed main effects of figural cut [$F(3, 141) = 28.25, p < .001$], ground cut [$F(3, 141) = 10.19, p < .001$], and their interaction [$F(9, 423) = 8.41, p < .001$], as shown in Fig. 7. Combinations that contained lighter and less saturated colors tended to be rated as more harmonious. The results of paired comparisons between each cut combination can be found in the Supplementary Material (Fig. S3), but to summarize: The L figures were judged most harmonious against all four ground cuts, and the D and S figures were judged least harmonious against all four ground cuts. These data are plotted as a function of hue difference in Supplementary Material Fig. S2, which shows that pair

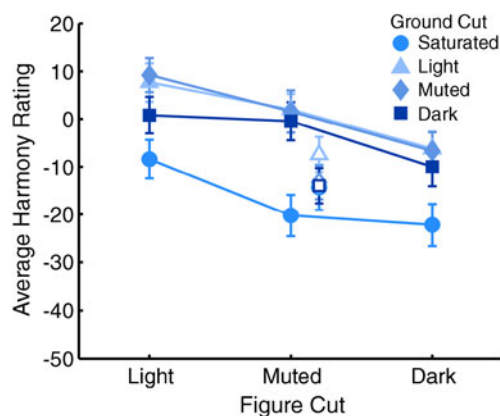


Fig. 7 Harmony ratings of color pairs for each ground cut (*separate lines*), as a function of figure cut (*x-axis*). Data points for the saturated figure cut (*open symbols*) are plotted separately at the same x-axis point as the muted colors because they share similar lightness levels, but they are slightly offset for clarity. Error bars standard errors of the means (SEM)

harmony decreased as hue difference increased for all cut-combinations, as they did for pair preferences in Experiment 1.

What, then, are the color appearance factors that influence ratings of color harmony? The same 10 Munsell factors tested for pair preference in Experiment 1 were analyzed in regression analyses to predict perceived color harmony. The best fitting model (Fig. 4) for the 992 color pairs showed that more harmonious pairs were more similar in hue, cooler, more desaturated, and more similar in coolness (67.3% of the variance explained). When all 10 Munsell factors are included in the full model, 72.6% of the variance could be explained, but there was no clear "best" model with more than four factors.

In the discussion of Experiment 1, we noted that one or more relational variables was required to account for the interaction between figure and ground colors in preference for color pairs. We then identified a set of relational Munsell factors that explained 53.5% of the variance in such preferences. When the inherently relational factor of harmony ratings is also included as a predictor variable, the best linear model accounts for 80.8% of the variance in preference ratings (multiple $r = +.90$). Harmony ratings alone explain 62.3% of the variance (more than all 10 Munsell factors combined), preference for the ground color adds another 9.3%, preference for the figure adds a further 4.7%, and the absolute value of the difference in Munsell values (lightnesses) adds a final 4.5% (larger lightness differences being preferred). Although there is a remarkably strong relation between harmony and preference, it falls considerably short of the equivalence that would be required to justify their interchangeable use by Chevreul (1839) and others (e.g., Granger, 1955a, 1955b, 1955c).

What are the differences between pair preference and harmony? Many are found in the effects of cuts (saturation

and lightness levels) where preferred pairs contain more dark and saturated colors and harmonious pairs are generally lighter (see Figs. 5 and 7). Figure 8 shows a scatter plot of preference ratings (y-axis) versus harmony ratings (x-axis) for each color pair in a way that highlights many of the principal differences. The high correlation between preference and harmony is evident in the strong linear trend of the point-cloud with a slope of somewhat less than unity. Differences between preference and harmony are then evident in systematic deviations from the best-fitting regression line.

First, Fig. 8 shows that the color pairs that are more preferred than harmonious (upper left quadrant) are generally high in lightness contrast, whereas those that are more harmonious than preferred (lower right quadrant) are generally low in lightness contrast. Second, it illustrates the dissociation between pair preference and pair harmony in terms of component color preferences. Palmer and Schloss (2010) found that the same participants especially disliked dark yellow and dark orange, and Fig. 8 shows that although pairs containing those particular colors were disliked, they were still judged to be harmonious when combined with light colors of similar hues. Figure 8 also highlights some similarities between preference and harmony. First, pairs containing cool colors are generally both more harmonious and more preferred (toward the upper right quadrant) than pairs containing warm colors, which are less harmonious and less preferred (toward the lower

left quadrant). Second, saturated red produces particularly disharmonious and disliked combinations (extreme lower left in Fig. 8), particularly those pairs containing a saturated red ground.

The differences between preference and harmony ratings can be analyzed quantitatively through regression analyses after their mutual variation (62.3%) has been removed. First, as stated above, the residual systematic variance in preference ratings was due to preferable ground colors (9.3%), preferable figural colors (4.7%), and large differences in lightness (4.5%). In contrast, the residual systematic variance in harmony ratings was due to greater hue similarity (i.e., fewer Munsell hue steps apart) (13.7%) and lower overall saturation (i.e., lower sum of the Munsell chroma coordinates) (6.4%). Altogether, pair preference, hue similarity, and low saturation explain 82.4% of the variance in average harmony ratings (multiple $r = +.91$). The latter two factors indicate that color pairs that are more harmonious than would be expected from preference ratings were the more similar pairs. Hue difference is clearly a similarity metric, but total saturation is also relevant to color similarity, because pairs of desaturated colors are closer to the center of color space, and all else being equal, closer together in color space than are highly saturated colors of corresponding hues and lightnesses.

We also examined spatial asymmetries in the pair harmony ratings due to figure-ground organization. The figure-ground asymmetry in lightness found for preference ratings ($r = +.14$) was also present in harmony ratings ($r = +.13$, $p < .001$), in that pairs with lighter figures on darker grounds were rated as more harmonious than pairs with darker figures on lighter grounds. However, the coolness asymmetry that was present in the preference ratings ($r = +.13$) failed to reach statistical significance in the harmony ratings ($r = +.05$, $p > .05$).

Although there is a high correlation between pair preference and pair harmony in the data averaged over all participants ($r = +.79$), the same is not necessarily true for individual participants. We computed each individual's degree of "preference-for-harmony" as the correlation between his/her preference ratings and his/her harmony ratings over all 992 color pairs. These correlations ranged from a high of 0.75, for the person who most preferred harmonious color combinations, to a low of -0.03 , for the person who was most indifferent to harmonious color combinations.³ We then examined a variety of factors that might predict these individual differences in preference-for-harmony, including their degree of training in color theory

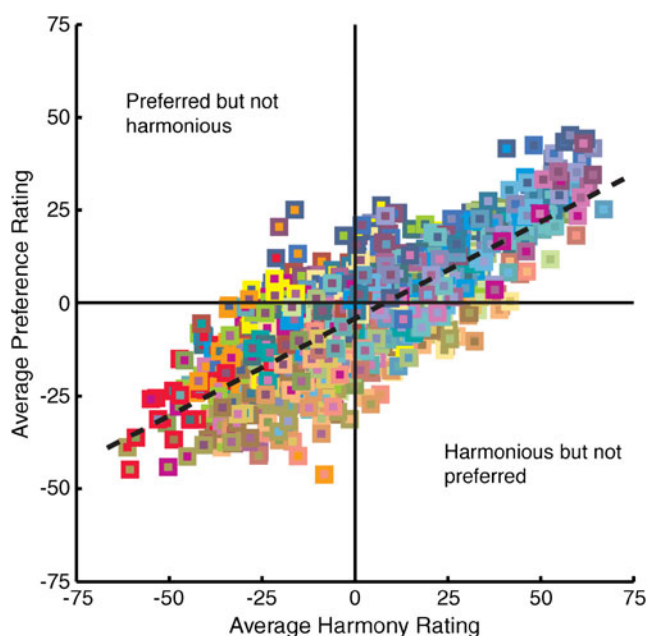


Fig. 8 Preference ratings for each color pair plotted as a function of its harmony rating. Each of the 992 data points depicts an approximation of the figural color (*small square*) and ground color (*large square behind the figure*). The *dashed line* shows the best fitting regression line relating preference to harmony ($y = -7.93 + 0.52x$)

³ It may initially seem odd that the highest individual correlation (+.75) is lower than the correlation of the group averages (+.79), but this only indicates that the pattern of deviations across individuals is noisy and tends to cancel out, on average, across individuals.

and the personality variables from the Big Five Inventory (or BFI) (John et al. 1991; John et al. 2008). (See Table S1 in the Supplementary Material for details.) The only factor that was reliably related to preference-for-harmony was the amount of formal color training that participants reported on a scale from 1 to 5, in response to the question, “How much formal training have you had in color?” Figure 9 shows average preference-for-harmony correlations plotted as a function of formal color training.

Somewhat surprisingly, preference-for-harmony was quadratically related to color training [$F(1, 47) = 7.58, p < .01$]. People who reported a moderate amount of formal training in color were most likely to prefer harmonious pairs. It is likely that everyone scoring 3 or above in color training was exposed to the kinds of rules that art theorists have formulated about color harmony and preference (e.g., Chevreul, 1839; Itten, 1973). Thus, they may well have been taught that harmonious combinations are preferable, and this pattern predominates among those with moderate color training. However, our participants who had more formal training, which included professional artists, decorators, and graphic designers, may have discovered through experience how to go beyond those rules in creating effective color combinations even with disharmonious pairs. Finally, those with essentially no formal training may simply have evaluated how much they like the two component colors in the pair, without much regard for the degree of harmony in those combinations. Supplementary Material Fig. S8 shows pair pair preferences separately for observers with low, moderate, and advanced formal training in color together with regression models that examine how perceived harmony and preference for component colors differentially explain pair preference ratings for each of the three groups.

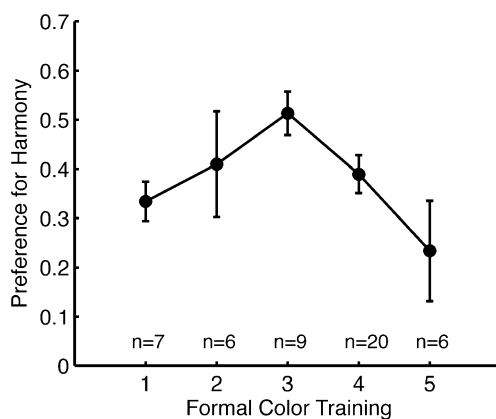


Fig. 9 Preference-for-harmony as a function of formal color training. Individual participants' correlations between their own pair preference ratings and pair harmony ratings are plotted as a function of level of formal color training, ranging from 1 = low to 5 = high. The number of participants in each group is displayed below the corresponding data point. Error bars standard errors of the means (SEM)

One question that can be asked about these harmony ratings is whether the instructions we gave produced a “demand characteristic” such that participants inferred that they are “supposed” to give the pattern of data that we observed. There are two noteworthy aspects of our instructions regarding color harmony. One is that they included the musical analogy, which explicitly told participants that their ratings of harmony did not need to conform to their ratings of preference. This analogy certainly does not dictate anything about how an individual “should” rate the harmony of a given color pair because the instructions specifically stated that “some [people] like Mozart and others like Stravinsky,” implying that harmony and preference ratings might be either quite similar or quite different. The other noteworthy aspect of the instructions is that they stated that harmonious colors are ones that “go naturally together.” Participants might have inferred from this that colors “should be” rated as harmonious to the extent that they are similar. This issue is addressed in Experiment 3, in which we obtain explicit ratings of color similarity and contrast them with ratings of harmony.

Experiment 3: Color similarity and its relation to preference and harmony

The results of Experiment 2 provided evidence that color harmony is not only closely related to color preference but also to color similarity: Harmonious colors are those with smaller hue differences, smaller differences in coolness, and lower total saturation, all of which imply that more harmonious colors are more similar to each other. We now address two further questions. First, how does color harmony differ from color similarity, if at all? Second, which of these two measures of color relations provides better predictions of pair preferences? If color harmony is, in effect, simply another name for color similarity, then similarity ratings should be able to explain as much variance in pair preferences as harmony ratings do. Moreover, there would be no need to consider the somewhat mysterious concept of color harmony if it predicts pair preference no better than the simpler concept of color similarity. Experiment 3, therefore, measures perceived color similarity of the same color pairs using the same BCP participants who previously made the preference and harmony ratings to examine more closely its relation to pair preference and pair harmony.

Methods

Participants The participants were the same 48 observers who completed Experiments 1 and 2.

Design and displays The design was the same as that of Experiments 1 and 2, but the displays were slightly different. The two colored regions were equal in size (100×100 pix) and positioned side by side, separated by a 20-pix gap. We did not use figure-ground displays for the similarity ratings because we wanted our observers to judge how similar the two component colors appeared to them without any spatial asymmetries in the displays (e.g., one color being inside another) or any complications arising from interactions along shared borders. Since all pair-wise combinations of the colors were tested, each pair appeared twice, once when one color appeared on the left and the other on the right, and a second time in the reversed spatial configuration. The left endpoint of the response scale was labeled “different” and the right endpoint was labeled “similar.”

Procedure As in Experiments 1 and 2, participants were presented with each of the 992 chromatic combinations one at a time in a random order. Their task was to rate how similar each pair of colors was on a scale from “different” to “similar.” Participants completed this task in a separate session, at least 1 day after the harmony task had been completed.

Results and discussion

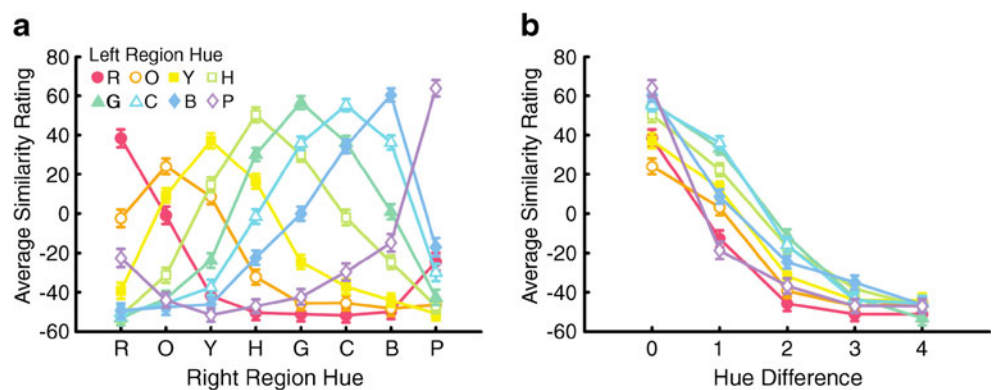
Average color similarity ratings are plotted in Fig. 10a as a function of figural hue and ground hue, averaged over figural cut and ground cut. As is evident by inspection, the hue effects on color similarity ratings are quite similar to the corresponding hue effects on harmony ratings plotted in Fig. 6a ($r = +.83$), but even more extreme. They are also somewhat similar to the preference ratings plotted in Fig. 2a ($r = +.55$). Color similarity ratings were highly consistent across subjects, with an average correlation of $+.75$ between each subject’s own ratings and the entire group’s average ratings. Notice that this consistency measure is

substantially greater than the same measure for both the harmony ratings [$r = +.51$, $t(47) = 8.84$, $p < .001$] and the preference ratings [$r = +.36$, $t(47) = 14.39$, $p < .001$].

The similarity data showed main effects of both figural hue [$F(7, 329) = 102.58$, $p < .001$] and ground hue [$F(7, 329) = 96.22$, $p < .001$], as well as a strong interaction between them [$F(49, 2303) = 174.77$, $p < .001$]. Like preference and harmony ratings, similarity ratings were highest for each pair when the figure and ground hues were the same and decreased as the hue difference increased. Figure 10b shows the same similarity data re-plotted as a function of the hue difference between the figure and ground colors. As was the case for the preference and harmony ratings in Figs. 2b and 6b, perceived similarity decreases monotonically as the hue difference between the two colors increases. The similarity functions do vary systematically over hue, however, with similarity being greater for the cool hues (blues, cyans, and greens) than for the warm colors (yellows, oranges, and reds) [$t(47) = 14.59$, $p < .001$], with purples and chartreuses being generally intermediate. The same data, plotted in terms of hue angles and differences in hue angles in CIELAB color space, can be found in the Supplementary Material (Figs. S12A and S12B, respectively).

Similarity ratings were also analyzed in terms of cut (saturation/lightness level). As shown in Fig. 11, there was a main effect of figure cut [$F(3, 141) = 52.13$, $p < .001$] ground cut [$F(3, 141) = 66.56$, $p < .001$], and a strong interaction between them [$F(9, 423) = 46.40$, $p < .001$]. Not surprisingly, pairs containing colors with more similar lightness values were rated as more similar. For example, dark colors were judged more similar to other dark colors than to muted colors [$t(47) = 4.53$, $p < .002$]. This pattern of results is different from color harmony ratings (Fig. 7), in which colors that generally contained lighter colors were more harmonious (e.g., dark colors were judged more harmonious with muted colors than with other dark colors [$t(47) = 4.26$, $p < .002$]). These data are plotted as a function of hue difference in Supplementary Material Fig. S4, which shows that, as for pair preference and pair harmony, similarity decreased as hue differences increased for all

Fig. 10 Similarity ratings for color pairs as a function of the hue on the right of the monitor (*x-axis*) and hue on the left of the monitor (*separate lines*) (a) and as a function of the hue difference (in terms of steps in the present BCP design) between the right and left colors (b). Error bars standard errors of the means (SEM)



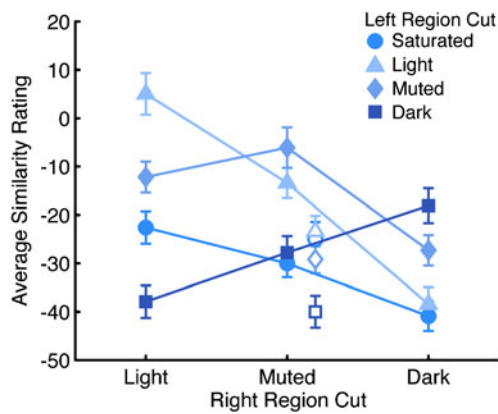


Fig. 11 Similarity ratings of color pairs for each left region cut (separate lines), as a function of right region cut (x -axis). Data points for the saturated figure cut (open symbols) are plotted separately at the same x -axis point as the muted colors because they share similar lightness levels but slightly offset for clarity. Error bars standard errors of the means (SEM)

combinations of cuts. Further analyses of the interaction between figure and ground cut as a function of hue difference between the two regions can be found in the Supplementary Material (Fig. S5).

When Munsell dimensions were used to predict color similarity ratings for the 992 color pairs, the best model showed that more similar colors were more similar in hue, cooler, more similar in value (lightness), and more similar in coolness, explaining 78% of the variance (see Fig. 4). When all 10 factors were included, the full model explained 82.8% of the variance, but there was no clear “best” model among those that included more than four predictors.

As noted previously, color similarity ratings are strongly correlated with harmony ratings ($r = +.83$). To analyze the differences between them, we looked at the residuals after removing their mutual variation (69.6%) through regression. The only additional predictor entered into the regression equation for harmony was the absolute value of the difference in Munsell value (+11.0%, with larger lightness differences being more harmonious) for a total of 80.6%, indicating that harmony ratings depended more strongly on lightness contrast (or less strongly on lightness similarity) than did similarity ratings. For similarity ratings, the absolute value of the difference in Munsell value explained an additional 12%, but unlike for harmony, smaller lightness differences were rated as more similar. An additional 7.5% of the variance can be explained by hue difference (the number of Munsell hue steps between the two colors), explaining a total of 89.1% of the variance.

The difference between perceived color similarity and color harmony, therefore, lies primarily in their relation to the lightness contrast of the two colors. Color similarity decreases as lightness contrast increases ($r = -.23$, $p < .001$, for the difference between the Munsell values/lightnesses of

the two colors), whereas harmony increases as lightness contrast increases ($r = +.10$, $p < .01$, for the corresponding difference). This pattern shows that our observers were not judging similarity when making their harmony ratings. If they were, the obtained dissociation between harmony and similarity in the lightness dimension would not be present. It also shows that our observers were not responding to a demand characteristic in which they inferred that harmony was the same as similarity, for their ratings clearly contradict this equivalence in the lightness dimension.

Thus far, we have established that color similarity is strongly related to, but not the same as, color harmony and that color harmony is strongly related to, but not the same as, preference for color pairs. This raises the important question of whether similarity is more useful in predicting pair preference than pair harmony is. The clear answer is: No. The raw correlation between average pair preference and average pair similarity ($r = +.55$) is substantially lower than the raw correlation between average pair preference and average pair harmony ($r = +.79$). A comparison between these correlations computed separately for each participant shows that the correlations between preference and harmony are reliably higher than those between preference and similarity [$t(47) = 8.24$, $p < .001$]. Indeed, if both average harmony ratings and average similarity ratings are included in the predictor variables of a regression analysis, similarity is never entered into the regression equation because it does not explain any additional variance in pair preference. If harmony ratings are not included, the best fitting regression model with similarity ratings accounts for 71.3% of the variance, substantially less than the 80.8% accounted for when harmony ratings are included.

Pair preference, harmony, and similarity are related to each other primarily because all of them increase as the hue similarity between the component colors increases: Color combinations with similar hues are generally more preferred, more harmonious, and more similar to each other. They differ primarily in terms of lightness contrast: Pair preference ratings depend more on lightness contrast than do harmony ratings, and harmony ratings depend more on lightness contrast than do similarity ratings.

Experiment 4: Preference for figural colors on background colors

Thus far, we have discussed preference and harmony judgments for color combinations as wholes and have found no evidence favoring art theoretic claims that color combinations with strong hue contrasts are either preferred or harmonious (e.g., Chevreul, 1839). One intriguing possibility is that the art theorists simply confused pair preference and pair harmony with what we are calling

figural preference. That is, people may find that figural colors are preferable against contrastingly colored backgrounds even though they do not find such pairs of colors either harmonious or preferred as combinations. This would be consistent with our previous finding that people prefer highly saturated colors to less saturated ones (Palmer & Schloss, 2010), because colors viewed against a background with a strongly contrasting hue are generally perceived as more saturated than when they are viewed against a background with a similar hue (e.g., Lotto & Purves, 2000). In **Experiment 4**, we studied how background color influences observers' preference for the figural color against which it was presented. We employed a rating task that was similar to Helson and Lansford's (1970) procedure to examine preferences for all 32 figural colors against all 32 background colors in an attempt to determine whether preferences for figural colors seen against different backgrounds vary in systematic ways that might explain art-theoretic claims about the aesthetic virtues of contrastive color combinations (e.g., Chevreul's so-called harmony of contrastive hues).

Methods

Participants The participants were the same 48 observers who completed **Experiments 1–3**. They performed the figural color-rating task on a different day that was later than the other three tasks.

Design and displays The eight colors from each of the four cuts were placed on each of the 32 background colors to make a total of 128 test displays, each containing all eight hues from the same cut on a uniform colored background. Each display contained the eight hues arranged to form a square with red in the top left corner, followed by orange, yellow, chartreuse, green, cyan, blue, and purple in a clockwise direction, as illustrated in Fig. 1a. Each colored square was 100×100 pix and was separated from the adjacent squares by 100 pix. In displays in which one of the squares was the same color as the background, that square was simply not visible in the display. Below each square was an asterisk, which marked the location of the response text box for each color. When participants typed in a rating, the asterisk below the colored square was replaced by the typed number.

The displays in **Experiment 4** (in which all eight colors from a given cut were presented simultaneously on a full-screen background color) were substantially different from the previous three experiments in which pairs were presented one at a time. We chose this configuration because we believed that it helped to emphasize that the task was to judge figural color preference independently of

the background color rather than preference for the figure ground combination as a whole.

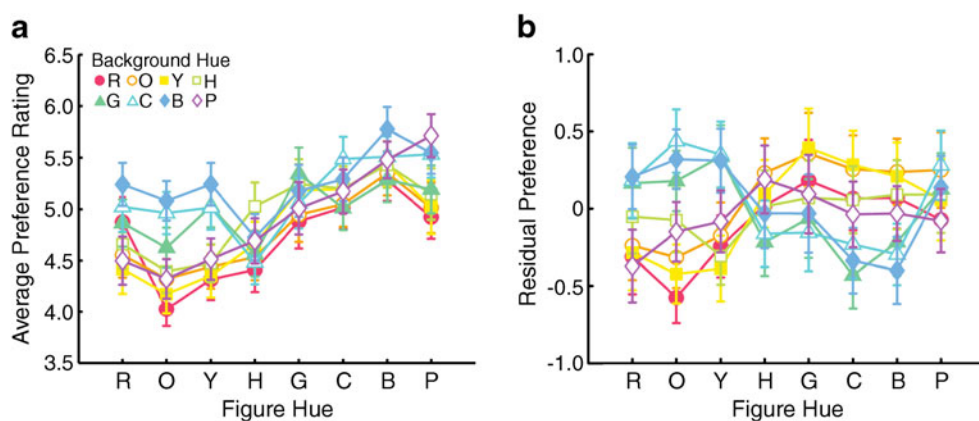
Procedure Each display contained the eight hues from one of the four cuts. Participants were asked to rate how much they liked each figural color on a scale from 1 (lowest) to 9 (highest) using the number keys at the top of the keyboard. They could rate the colors in any order they wished, using the tab key to select which colored square to rate. When a square was selected, the asterisk below it enlarged so that participants knew which square they were currently expected to rate. If they desired, participants could change their ratings by tabbing back to a color square and typing a new rating. In displays that contained a figural color that was identical to the ground, there was a zero below the square instead of the asterisk, and that square was skipped when the tab key was pressed.

Participants were told that a given color could look different on different backgrounds, so they need not try to be consistent in their ratings across trials. In addition, they were informed that they could give multiple colors the same rating within a given trial (i.e., if they hated all the colors they could give them all a rating of "1" and if they loved them all they could give them all a rating of "9." Once participants had rated all the colors in a test display, they pressed the "Enter" key to go onto the next display. The 128 displays were presented in a random order and were separated by a 500-ms inter-trial interval.

Results and discussion

Figure 12a plots the preferences for figural hues on different colored backgrounds as a function of background hue. This pattern is somewhat similar to the pair preferences presented in Fig. 2a ($r = +.54$) but is also clearly quite different in that the ground color curves do not peak when the figural color has the same hue, as they do in Fig. 2a. When figural preferences for each of the 32 figural colors (averaged over backgrounds) were compared with pair preferences for the same figural colors within figure-ground pairs (also averaged over backgrounds), there was a strong correlation ($r = +.74$), but it was not as strong as preferences for the same 32 figural colors when viewed against a neutral gray background (Palmer & Schloss, 2010) ($r = +.87$). Indeed, when these two correlations are calculated separately for each individual participant and compared statistically, correlations between figural color preference on differently colored backgrounds were reliably more closely related to figural color preferences on a neutral gray background than to pair preferences in which that color is figural [$t(47) = 4.64$,

Fig. 12 Preference ratings for each figural hue on each of the background hues as a function of figural hue (a) and residual figural color preference after accounting for figural preferences when rated on a neutral gray background (Palmer & Schloss, 2010) and pair preferences plotted as a function of figural hue (b). Error bars standard errors of the means (SEM)



$p < .001$]. This finding strongly suggests that the observers in Experiment 4 were indeed rating how much they preferred the figural colors in the present task rather than how much they liked the figure-ground pairs as wholes. The same data, plotted as a function of figural hue angle in CIELAB color space, can be found in the Supplementary Material (Fig. S13A).

There was a main effect of figural hue [$F(7, 329) = 7.70$, $p < .001$] and ground hue [$F(7, 329) = 8.47$], and an interaction between them [$F(49, 2,303) = 4.58$, $p < .001$] indicating that figural color preferences are indeed influenced by ground color. As is evident in Fig. 12a, figural colors were more preferred on cooler backgrounds [$t(47) = 5.27$, $p < .001$]. This was especially true for the warm figural colors (red, orange, and yellow) against the cool backgrounds (blue, cyan, and green) compared with warm figural colors against warm backgrounds [$t(47) = 6.03$, $p < .001$].

A regression model was used to predict preference for figural colors on different colored backgrounds using the same ten Munsell factors as predictors (see Experiments 1, 2 and 3). The best model (Fig. 4) showed that figural colors were more preferred when they contrasted with the background lightness/value, when they and the background were cooler, when they were more saturated and cooler than the background, and when they and the background were more saturated (58.4% of the variance explained). A total of 62.3% was explained when all 10 factors were included in the full model, but there was no clear “best” model containing more than 5 factors.

When single-color preferences for the figural color and the ground color (each rated independently by the same observers against a neutral gray background color; see Palmer & Schloss, 2010), pair preferences, pair harmonies, and pair similarities were included in a regression model together with the Munsell factors, a total of 66.0% of the variance in figural preference against colored backgrounds was explained by color preference for the figural color on a gray background (30.3%), pair preference (18.7%), pair similarity (12.0%; larger differences being more preferred),

and signed chroma/saturation difference (5.0%; more saturated figures on more desaturated grounds being preferred). The increase in figural color preference as perceived similarity decreases is the first evidence we have obtained that preference of any sort increases as hue contrast increases.

To look more closely at the effects of hue contrast, Fig. 12b plots the residual figure preferences after removing the variance due to other sorts of preference: namely, preference for the figural color when viewed against a neutral gray background and preferences for pairs containing the relevant color as figure. There is a clear interaction in the residuals in which warmer hues are preferred on cooler backgrounds and cooler hues are preferred on warmer backgrounds [$F(49, 2,303) = 7.69$, $p < .001$]. This pattern is clearer for the “core” cool hues (green, cyan, and blue) and the “core” warm hues (red, orange, and yellow), than for the “border” hues (chartreuse and purple). Chartreuse followed a similar pattern to the warm hues, but purple peaked over chartreuse, which is the hue that contrasts most with purple. The residual figural color preferences are plotted as a function of figural hue angle (CIELAB) in the Supplementary Material (Fig. S13B).

An analysis of the effects of cuts on figural color preference showed a main effect of figural cut [$F(3, 141) = 8.16$, $p < .001$], ground cut [$F(3, 141) = 8.77$, $p < .001$], and a strong interaction between them [$F(9, 423) = 20.84$, $p < .001$]. As shown in Fig. 13, saturated figures are generally most preferred, colors on saturated grounds are generally least preferred, light figures are more preferred on dark backgrounds, dark figures are more preferred on light backgrounds, and colors are moderately preferred on muted backgrounds (see Supplementary Material Fig. S7 for supporting statistics of all pairwise comparisons).

Detailed discussion of evidence for figural preference increasing as both hue and lightness similarity decreases can be found in the Supplementary Material (Figs. S6 and S7). In summary, preference for figural colors, combined across hue and cut, increased as hue difference between the

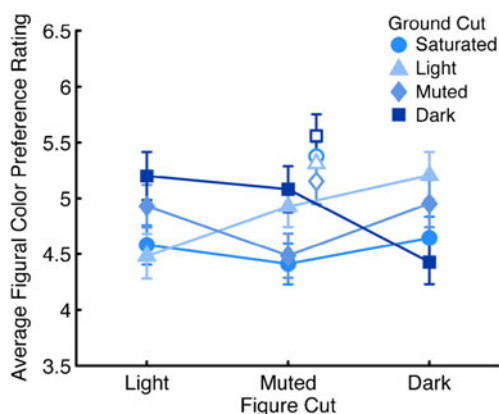


Fig. 13 Preference for figural cuts (*x*-axis) on different background cuts (*separate lines*). Data points for the saturated figure cut (*open symbols*) are plotted separately at the same *x*-axis point as the muted colors because they share similar lightness levels, but they are slightly offset for clarity. *Error bars* standard errors of the means (SEM)

figure and background increased, which is the opposite of the pattern for pair preference, harmony, and similarity. Upon a closer examination, this pattern is primarily limited to color pairs with similar lightness levels, which suggests that hue contrast is more preferable only when there is minimal lightness contrast.

The results of this experiment are roughly consistent with art theorists' claim that hue contrast enhances people's preference for colors in combinations that contain at least certain kinds of hue contrast (e.g., Chevreul, 1839; Munsell, 1921). The main problem with the art theoretic claims is that it is misattributed to increased harmony. In fact, people do not like strong hue contrasts because such combinations are *harmonious*; they like colors against strongly contrastive backgrounds because they make the figural color itself look "better" (more preferred) than it does against a weakly contrastive background. This argument is consistent with the fact that people generally prefer saturated colors over the other three less-saturated cuts when rated on a (zero saturation) neutral gray background (Palmer and Schloss 2010): saturated colors are more contrastive than other colors against medium gray. We will discuss why people might prefer colors against strongly contrastive backgrounds in the "General discussion" section where we address the general question of possible causes of the effects reported in this article.

Thus, it appears that virtually all of the residual effects in these figural color preferences, after variations due to single and pair preferences have been removed, can be attributed to some form of contrast, all of which generally enhance preference for the figural color. In summary, the results show that figural color preference increases as hue similarity decreases, which is opposite the pattern for pair preference ratings, harmony ratings, and similarity ratings

obtained in Experiments 1–3, respectively. They also show that pairs are most preferred on backgrounds of contrasting lightness.

Our results thus generally support Helson and Lansford's (1970) claim that contrast is a highly influential factor on how much people like figural colors (which they call "object colors") against a background color. They propose that the reason contrast improves figural color preference could be ease of perception on a contrasting background. This fits with the idea that preference in general is related to perceptual "fluency:" the hypothesis that people aesthetically prefer displays that are easier to perceive (e.g., Reber, Schwarz, & Winkielman, 2004).

Finally, we performed a regression analysis to predict pair preference (Experiment 1) from figural color preference (Experiment 4), as well as pair harmony ratings (Experiment 2), similarity ratings (Experiment 3) and Munsell factors. The best-fitting model, which explained 82.6% of the variance in pair preferences, included harmony (62.3%), figural color preference when rated on the correspondingly colored background (12.3%), and ground color preference on a neutral gray background (+8%). This amount is only slightly more than the model from Experiment 2 (80.8%) that included figural color preference on a neutral gray background and lightness contrast, both of which are encapsulated by figural color preference on different colored backgrounds. Nevertheless, this model, which accounts for the most variance with the fewest variables, supports the hypothesis that contextual preference for the figural color (i.e., figural preference on a colored background) has an effect on pair preferences, even though it does not have an effect on pair harmony.

General discussion

In this article, we have shown that there are distinct differences among three kinds of perceptual judgments of two-color figure-ground combinations: preference for the color pair, harmony of the color pair, and preference for figural colors against colored backgrounds. Both pair preference and pair harmony vary primarily as a function of hue similarity, such that pairs with similar hues are, on average, both more preferred and more harmonious. Consistent with color theories in art (e.g., Chevreul's 1839 "harmony of analogous colors"), ratings of color preference and harmony were highest for colors most similar in hue. Inconsistent with such theories (e.g., Chevreul's "harmony of contrastive colors"), however, no overall increase was observed in ratings of preference or harmony for complementary hues.

Although preference and harmony are closely related to one another, preferred pairs differ from harmonious pairs in including preference for the component colors and a large

lightness contrast component, whereas harmonious pairs are more similar in hue and lower in saturation. Harmony and similarity ratings are also closely related to one another, but harmony ratings do not have the lightness similarity component that similarity ratings have.

Finally, figural color preferences against different background colors are closely related to preference for the same figural colors when rated on a neutral gray background and preference for the combination of the figural color and background color. Once those factors are accounted for, however, clear effects of both hue contrast and lightness contrast are revealed: warmer figures are preferred on cooler backgrounds, cooler figures are preferred on warmer backgrounds, and figures are generally preferred on backgrounds of contrasting lightness. These results show that Chevreul's so-called "harmony of contrast," at least in the hue dimension, actually applies to preferences for figural colors on different colored backgrounds rather than to pair preferences or pair harmonies.

The present experiments were aimed primarily at establishing the nature of aesthetic preferences for color pairs and their relations to harmony, similarity, and figural preference of color pairs. From these data, we can infer little about the actual causes of pair preferences. Still, we can speculate about causes with varying degrees of confidence for several key aspects of our findings. The primary factors that influence pair preferences appear to be preferences for single colors (of the individual figural color and/or ground color), color harmony of figure and ground, lightness contrast between figure and ground, and figural preference against a colored ground. Before closing, we will consider in turn what factors might underlie each of these factors.

The data from [Experiment 1](#) clearly show that people's preferences for color pairs reliably depend on their preferences for the individual colors of which they are composed (e.g., see [Fig. 3](#)). [Palmer and Schloss \(2010\)](#) have reported results that strongly support an ecological valence theory (EVT) of single color preferences, positing that people like colors to the degree that they like correspondingly colored objects. For example, people generally like saturated blues and cyans because they like clear sky, clean water, swimming pools, and most other objects that characteristically are these colors. They generally dislike dark oranges (browns) and dark yellows (olive-colors) because they dislike feces, rotting food, vomit, and many other (but not all other, consider chocolate and coffee) objects they associate with these colors. Because one cannot make scientific generalizations about such observations based on just a few examples of desirable and undesirable colored objects, [Palmer and Schloss](#) devised a systematic procedure to test their theory.

To obtain comprehensive lists of color–object associations, one group of participants provided verbal descriptions of all the objects they associated with each of the 32 BCP colors in a fixed time period. Another group then rated their affective valence for each verbally described object (i.e., how positive/negative they felt about "clear sky," "feces," etc.). A third group rated how well the colors of each verbally described object matched the BCP color(s) that had elicited it. The affective valence ratings for each described object were weighted (multiplied) by the relevant color-match ratings (higher match ratings produced higher weights) and then averaged for each of the 32 BCP colors to produce the weighted affective valence estimate (WAVE) for each color. The WAVE for a given color, therefore, was calculated as the average weighted valences of *all* objects associated with that color, which could range from very positive to very negative. For example, object associates for brown (BCP dark orange) included "chocolate," which was very positive, "feces," which was very negative, and a large number of other objects with intermediate valences, all of which averaged together gave a net negative WAVE for this color). Using this procedure for all 32 chromatic colors, [Palmer and Schloss \(2010\)](#) found a strong correlation between the WAVEs of the BCP 32 colors and people's average preference ratings for the same 32 colors ($r = +.89$). This result shows that preference for a given color increases as the average weighted valence of all of the objects associated with that color increases.

Because the same single color preferences appear in the regression models for the present pair preferences (see also [Fig. 3a](#) for pair preference ratings averaged over ground hue and figure hue versus [Fig. 3b](#) for single color preferences), we assume that this component of the data from [Experiment 1](#) is influenced by the same ecological valences. Moreover, to the extent that certain color combinations are characteristic of entities with strong valences (e.g., red and green with Christmas, blue and yellow with a bright sun against a clear sky, and dark purple and dark green with bruised flesh), the same associative ecological valence principles suggest that color pairs may be more (or less) preferred than would otherwise be expected from the kind of colorimetric relations we have identified in this article (e.g., hue similarity and lightness contrast), depending on the valence of their ecological associations. Of course, one cannot simply point to cherry-picked examples of objects that are associated with color pairs to test for ecological effects on pair preferences. A comprehensive analysis of *all* objects (positive, negative, and everything in between) associated with each color pair would be necessary to test whether the average valence of objects associated with a given color pair is related to preference for that same pair.

The lion's share of the variance in pair preferences, however, is clearly due to abstract color relationships: people prefer color pairs that have the same hue but differ in lightness and/or saturation. Our measurements suggest that the best single relational variable in predicting pair preferences is perceived pair harmony, because average pair harmony ratings appear in all of the best-fitting regression models of average pair preference, accounting for 62% of the variance. What, then, might be the cause of the perception of color harmony? Our instructions for rating harmony (aside from the musical analogy) asked observers to report "how well the colors go together," and we presume that this is what they judged, to the best of their ability. Our current conjecture is that color harmony derives from the ecological co-occurrence statistics of color pairs within uniform connected (UC) regions of natural images. Palmer and Rock (1994) defined UC regions as connected areas within an image that are (relatively) homogeneous in terms of many variables, including those related to color. We speculate that the color pairs judged to be most harmonious are those that are most likely to co-occur within UC regions. We are testing this hypothesis by examining ecological statistics in the Berkeley Segmentation Dataset (see Martin, Fowlkes, Tal, & Malik, 2001; <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>), which contains 200 images that were hand-parsed into regions by human observers. Preliminary results from analyses of the relations among within-region colors suggest that the primary chromatic attribute defining a UC region is hue similarity. This means that pairs of pixels that have the same hue (or very similar hues) are most likely to co-occur within UC regions and that within-region variations in lightness and/or saturation are greater than variations in hue.

A third factor that clearly contributes to pair preference in most of the best-fitting regression models is lightness contrast. Harmony ratings do not depend strongly on lightness contrast, but pair preferences do, with more contrastive pairs being preferred. Why might this occur? One possible explanation comes from the fluency theory of aesthetic preference (e.g., Reber et al., 2004). The basic premise of fluency theory is that people prefer things that are easy to process perceptually. Lightness contrast is one of the primary factors that supports this theory: People prefer images in which the contrast between figure and ground regions is high. Fluency theory frames the relevance of lightness contrast to pair preference in terms of high-contrast figure-ground images being aesthetically pleasing to perceive, but one could also frame the same phenomenon in the opposite terms: Perhaps low-contrast figure-ground images are aesthetically displeasing. This description suggests a possibly different causal account in which isoluminance plays a dominant role: Perhaps people dislike low contrast figure-ground displays as the colors approach

isoluminance, making the boundaries between them difficult to discriminate and having perceptually disturbing effects (e.g., Gregory, 1977). We are currently investigating these possibilities, both of which may contain some truth.

Finally, the results of **Experiment 4** suggest that hue contrast increases preference for a figural color against a colored background. This effect may be caused by simultaneous color contrast (also known as induced color). The background (or surround) induces a hue shift in the figural color that is complementary to the background color (e.g., da Vinci, 1492; Chevreul, 1839; Helmholtz, 1866/1925; Walraven, 1976; Shevell, 1978). This means that a gray figure on a blue background should appear somewhat yellowish (because yellow is the complement of blue), a yellow figure on a blue background should appear extra yellow (because the yellowness induced by the blue background increases the saturation of the yellow figure), and a blue figure on blue background should appear somewhat grayish (because the yellowness induced by the blue background partly cancels the blueness of the figure). If people generally like more saturated figural colors, as they apparently do (Palmer & Schloss, 2010), and if a contrasting background enhances the saturation of the figural color, then figural colors should be more preferred on backgrounds with strongly contrastive hues. The key question is whether these hue contrast effects will be eliminated if observers first adjust each color on each colored background to look identical to that same color on a neutral gray background. If all the figural preference effects found in **Experiment 4** were to disappear with the appearance-matched figural colors, then simultaneous color contrast is surely their cause. We are currently investigating this possibility.

Before closing, we want to briefly address two concerns that can be raised about the generalizability of the current results. The first concerns the degree to which preferences for color pairs in concentric-square, figure-ground displays will generalize to preferences for color pairs displayed in other spatial configurations. Preliminary data for color pairs displayed side-by-side with a gap between them suggest that pair preferences still generally increase as hue similarity between the component colors increases. Naturally, certain kinds of spatial factors that produce semantic interpretations of the colored regions could produce fairly pronounced effects on preferences, such as making an orange region carrot-shaped and a green region above it carrot-top-shaped. Future work comparing the influence of different geometric arrays is underway to test how the principles established in this article apply to color combinations in different spatial arrangements.

The second issue concerns how the present findings from two-color combinations might generalize to combinations of

three colors, four colors, and beyond. The present data show that preferences for single component colors only weakly predict preference for color pairs, with the lion's share of the variance attributable to pairwise color relations (e.g., harmony). Might the same problem arise when expanding the domain to three-color combinations: i.e., might single and pairwise preferences account for little of the variance, with the lion's share now arising from three-way relationships? Preliminary results on preference for color triples, however, suggest that preferences for all possible pairs within triples of colors predict much of the variance within preference for triples as a whole. We speculate, therefore, that once pairwise color preferences are known and understood, enough relational information is available to account for preferences in higher-order combinations.

At the outset of this paper, we proposed that much of the confusion in the literature on the aesthetics of color combinations was due to confusion among three distinct types of judgments: pair preference, pair harmony, and figural preference on different colored backgrounds. We have provided strong empirical evidence that these three types of judgments are indeed different, in that they produce systematically different patterns of results. We have also argued that our results and analyses clarify many of the confusions that have accumulated over the past century. Moreover, we expect that the new understanding achieved by making clear distinctions among these and related aspects of perceptual response will allow researchers to move beyond the foundational problems of how to define and measure preference and harmony properly to more advanced questions, such as *why* people prefer the combinations they do, both as individuals and as a group, and how color preferences might be influenced by the context and/or intended message of a visual display. We believe that the time is ripe to answer such questions.

Acknowledgments We wish to thank Jeremy Wolfe, David Simmons, and Stephen Westland for their helpful comments on this manuscript. We also thank Rosa Poggesi, Christine Nothelfer, Patrick Lawler, Laila Kahn, Cat Stone, Divya Ahuja, and Jing Zhang for assistance with the collection of data reported in this article and Joseph Austerweil for help with statistical analyses. The material described in this article is based on work supported by the National Science Foundation under Grant No. 0745820 and a Google Gift to S.E.P., as well as by a generous gift of product coupons from Amy's Natural Frozen Foods (Santa Rosa, CA) to K.B.S., with which we "paid" many of our participants. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, Google, or Amy's Natural Frozen Foods.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix

Table 1 CIE 1931 values and Munsell values for the 32 BCP colors. The first letter in the color name refers the cut (saturated (S), light (L), muted (M), and dark (D)). The second letter in the name refers to the hue (red (R), orange (O), yellow (Y), chartreuse (H), green (G), cyan (C), blue (B) and purple (P))

Color	x	y	Y (cd/m ²)	Hue	Value	Chroma
SR	0.549	0.313	22.93	5 R	5	15
LR	0.407	0.326	49.95	5 R	7	8
MR	0.441	0.324	22.93	5 R	5	8
DR	0.506	0.311	7.60	5 R	3	8
SO	0.513	0.412	49.95	5 YR	7	13
LO	0.399	0.366	68.56	5 YR	8	6
MO	0.423	0.375	34.86	5 YR	6	6
DO	0.481	0.388	10.76	5 YR	3.5	6
SY	0.446	0.472	91.25	5 Y	9	12
LY	0.391	0.413	91.25	5 Y	9	6.5
MY	0.407	0.426	49.95	5 Y	7	6.5
DY	0.437	0.450	18.43	5 Y	5	6.5
SH	0.387	0.504	68.56	5 GY	8	11
LH	0.357	0.420	79.90	5 GY	8.5	6
MH	0.360	0.436	42.40	5 GY	6.5	6
DH	0.369	0.473	18.43	5 GY	4.5	6
SG	0.254	0.449	42.40	3.75 G	6.5	11.5
LG	0.288	0.381	63.90	3.75 G	7.75	6.25
MG	0.281	0.392	34.86	3.75 G	6	6.25
DG	0.261	0.419	12.34	3.75 G	3.75	6.25
SC	0.226	0.335	49.95	5 BG	7	9
LC	0.267	0.330	68.56	5 BG	8	5
MC	0.254	0.328	34.86	5 BG	6	5
DC	0.233	0.324	13.92	5 BG	4	5
SB	0.200	0.230	34.86	10 B	6	10
LB	0.255	0.278	59.25	10 B	7.5	5.5
MB	0.241	0.265	28.90	10 B	5.5	5.5
DB	0.212	0.236	10.76	10 B	3.5	5.5
SP	0.272	0.156	18.43	5 P	4.5	17
LP	0.290	0.242	49.95	5 P	7	9
MP	0.287	0.222	22.93	5 P	5	9
DP	0.280	0.181	7.60	5 P	3	9

References

- Albers, J. (1971). *Interaction of color*. New Haven: Yale University Press.

- Allen, E. C., & Guilford, J. P. (1936). Minor studies from the psychological laboratory of the University of Nebraska X Determining the affective values of color combinations. *The American Journal of Psychology*, *48*, 643–648. doi:10.2307/1416516
- Burchett, K. E. (2002). Color harmony. *Color Research and Application*, *27*, 28–31. doi:10.1002/col.10004
- Camgöz, N., Yerner, C., & Güvenç, D. (2002). Effects of hue, saturation, and brightness on preference. *Color Research and Application*, *27*, 199–207. doi:10.1002/col.10051
- Chevreul, M. E. (1839). *The principles of harmony and contrast of colors*. Birren, F. (Ed.) (1987). New York: Van Nostrand Reinhold.
- Chuang, M., & Ou, L. (2001). Influence of a holistic color interval on color harmony. *Color Research and Application*, *26*, 29–39.
- da Vinci, L. (1492). *Treatise on painting*. (A. P. McMahon, 1956, Trans). Princeton: Princeton University Press.
- Goethe, J. W. (1810/2006). *Theory of colours*. Translated by C. L. Eastlake (1840) from German edition “Farbenlehre.” Mineola: Dover Publications.
- Granger, G. W. (1952). Objectivity of colour preferences. *Nature*, *170*, 228–280.
- Granger, G. W. (1953). Area balance in color harmony: An experimental study. *Science*, *117*, 59–61. doi:10.1126/science.117.3029.59
- Granger, G. W. (1955a). An experimental study of colour harmony. *Journal of General Psychology*, *52*, 21–35.
- Granger, G. W. (1955b). Aesthetic measure applied to color harmony: An experimental test. *Journal of General Psychology*, *52*, 205–212.
- Granger, G. W. (1955c). The prediction of preference for color combinations. *Journal of General Psychology*, *52*, 213–222.
- Gregory, R. L. (1977). Vision with isoluminant colour contrast: 1. A projection technique and observations. *Perception*, *6*, 113–119.
- Hård, A., & Sivik, L. (1981). NCS - Natural Color System: A Swedish standard for color notation. *Color Res & Appl*, *6*, 129–138.
- Helson, H., & Lansford, T. (1970). The role of spectral energy of source and background color in the pleasantness of object colors. *Applied Optics*, *9*, 1513–1562.
- Helmholtz, H. von (1866/1925). In J. P. C. Southall (Ed.), *Treatise on physiological optics* (Translated from the 3rd German edition). New York: Dover Publications.
- Hurlbert, A. C., & Ling, Y. (2007). Biological components of sex differences in colour preferences. *Current Biology*, *17*, 623–625.
- Itten, J. (1973). *The art of color*. New York: Van Nostrand Reinhold. Originally printed in 1961.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory—versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big-five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York: Guilford Press.
- Lotto, R. B., & Purves, D. (2000). An empirical explanation of color contrast. *Proceedings of the National Academy of Science*, *97*, 12834–12839.
- Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). *A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics*. Paper presented at the 8th International Conference on Computer Vision, July 7–14, Vancouver, British Columbia.
- Moon, P., & Spencer, D. E. (1944a). Geometric formulation of classical color harmony. *Journal of the Optical Society of America*, *34*, 46–59.
- Moon, P., & Spencer, D. E. (1944b). Aesthetic measure applied to color harmony. *Journal of the Optical Society of America*, *34*, 234–242.
- Munsell, A. H. (1921). *A grammar of color*. New York: Van Nostrand Reinhold.
- Munsell, A. H. (1966). *The munsell book of color-glossy finish collection*. Baltimore: Munsell Color Company.
- Nemcsics, A. (1993). *Colour dynamics: Environmental colour design*. New York: Ellis Horwood.
- Nemcsics, A. (2007). Experimental determination of laws of color harmony. Part 1: Harmony content of different scales with similar hue. *Color Research and Application*, *32*, 477–488.
- Nemcsics, A. (2008). Experimental determination of laws of color harmony. Part 2: Harmony content of different monochrome pairs. *Color Research and Application*, *33*, 262–270. doi:10.1002/col.20416
- Nemcsics, A. (2009a). Experimental determination of laws of color harmony. Part 3: Harmony content of different hue pairs. *Color Research and Application*, *34*, 33–44. doi:10.1002/col.20457
- Nemcsics, A. (2009b). Experimental determination of laws of color harmony. Part 4: Color preference and the color harmony content. *Color Research and Application*, *34*, 210–224. doi:10.1002/col.20489
- Ostwald, W. (1932). *Color science* (Vol. 1). London: Windsor Newton.
- Ou, L., & Luo, M. R. (2006). A colour harmony model for two-colour combinations. *Color Research and Application*, *31*(3), 191–204. doi:10.1002/col.20208
- Ou, L., Luo, M. R., Woodcock, A., & Wright, A. (2004a). A study of colour emotion and colour preference, part II: Colour emotions for two-color combinations. *Color Research and Application*, *29*, 292–298. doi:10.1002/col.20024
- Ou, L., Luo, M. R., Woodcock, A., & Wright, A. (2004b). A study of colour emotion and colour preference, part III: Colour preference modeling. *Color Research and Application*, *29*, 381–389.
- Palmer, S. E., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review*, *1*(1), 29–55.
- Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Science*, *107*(19), 8877–8882.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the Perceiver’s processing experience? *Personality and Social Psychology Review*, *8*, 364–382.
- Schloss, K. B., & Palmer, S. E. (2010). *The role of spatial organization in preference for color pairs*. Manuscript in preparation.
- Schloss, K. B., Poggesi, R. M., & Palmer, S. E. (2010). *Effects of university affiliation and “school spirit” on color preferences: Berkeley vs. Stanford*. Manuscript submitted for publication.
- Shevell, S. K. (1978). The dual role of chromatic backgrounds in color perception. *Vision Research*, *18*, 1649–1661.
- Walraven, J. (1976). Discounting the background—the missing link in the explanation of chromatic induction. *Vision Research*, *16*, 289–295.
- Westland, S., Laycock, K., Cheung, V., Henry, P., & Mahyar, F. (2007). Colour harmony. *Colour: Design & Creativity*, *1*, 1–15.
- Wyszecki, G., & Stiles, W. S. (1967). *Color science: Concepts and methods, quantitative data and formulas*. New York: John Wiley.