Aesthetic Response to Color Combinations: Preference, Harmony, and Similarity

Supplementary Material

Karen B. Schloss and Stephen E. Palmer

University of California, Berkeley

Effects of Cut on Pair Preference, Pair Harmony, Pair Similarity, and Figural Color Preference.

To understand the nature of pair preferences, pair harmony, pair similarity, and figural color preference beyond the findings described in the main article, we examined the effects of the component colors' cuts through color space – saturated (S), light (L), muted (M) and dark (D) – as a function of the hue difference between the two colors. There were no data points for same-cut pairs with zero hue difference because the component colors were identical, so we conducted separate ANOVAs on two data sets. The first ANOVA included only hue-difference steps of 1 through 4 for all cut comparisons (hereafter "Δ1-4, all cuts"). The second ANOVA included all five hue-difference steps (0 through 4) but only for cut comparisons in which the figure and ground were from different cuts (hereafter "Δ0-4, different cuts"). The y-axis limits in Figures S1 (pair preference), S2 (harmony), and S4 (similarity) are the same to allow direct comparisons between the figures. The scale for figural color preferences in Figure S6 is different, however, because those data were collected on a different ratings scale (1-9) than that the other three measures (-100 to 100).

As shown in Figure S1, pair preferences for all cut combinations decreased monotonically as hue difference between the figure and ground color increased (Δ1-4, all cuts: $F(3,141) = 27.83$, $p < .001$); Δ0-4, different cuts: $F(4,188) = 51.08$, $p < .001$). There was a 3-way interaction among figure cut, ground cut, and hue difference (Δ1-4, all cuts: $F(27,1269) = 2.22$, $p < .001$; Δ0-4, different cuts: $F(24, 1128) = 4.31$, $p < .001$), but the size of this interaction is small, its nature is unsystematic, and its interpretation is unclear. The simpler figure cut x hue difference interaction was systematic, however, with preference for pairs with D and M figures decreasing more rapidly than pairs with S and L figures as hue difference increased (Δ1-4, all cuts: $F(9, 423) = 2.25$, $p < .05$; Δ0-4, different cuts: $F(8, 376) = 4.29$, $p < .001$). There was no such interaction between ground cut and hue difference (Δ1-4, all cuts: $F(9, 423) = 1.16$, $p > .05$; Δ0-4, different cuts: $F(12, 564) = 1.61$, $p > .05$).

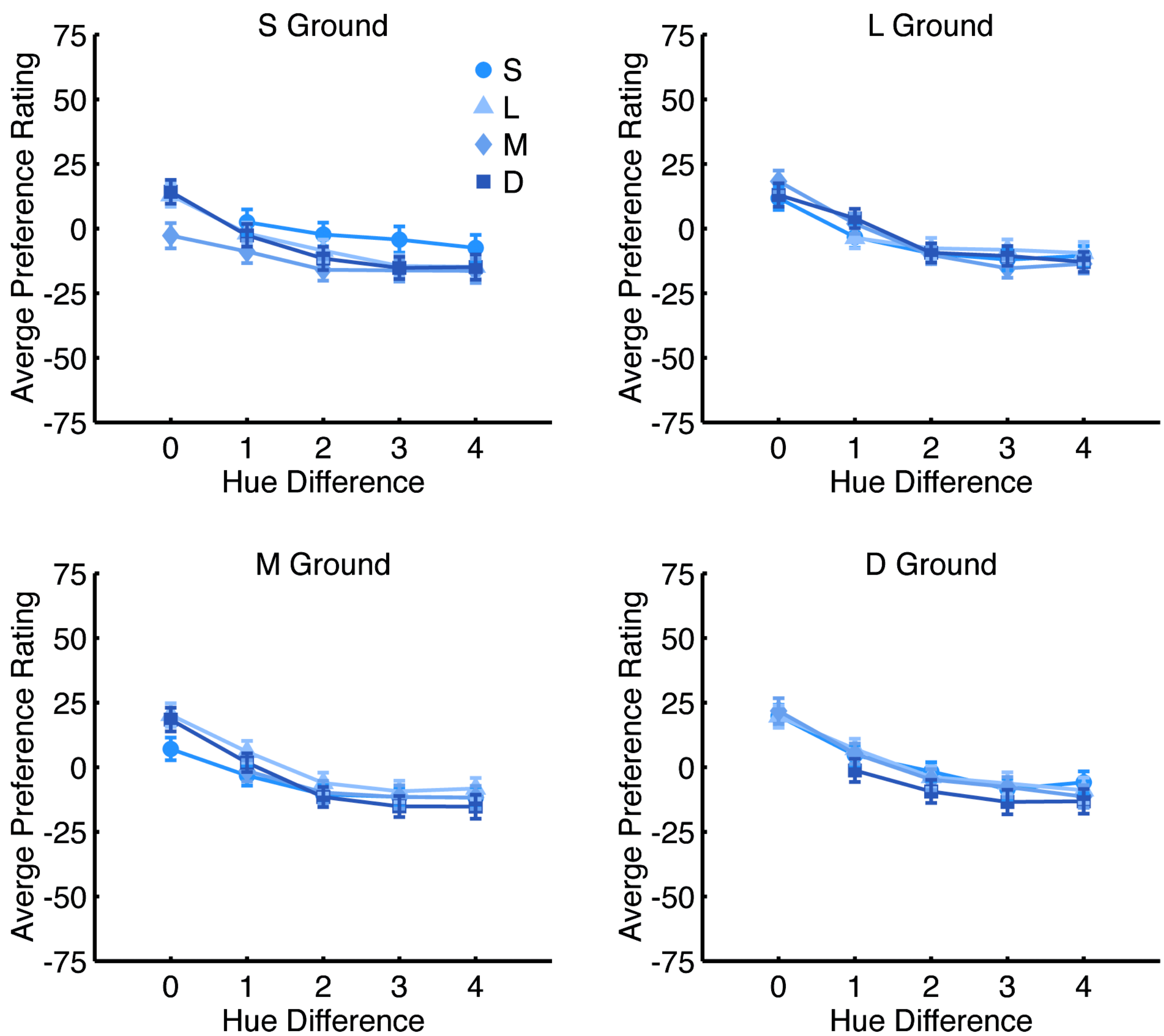

Figure S1. Pair preference ratings for figural color (separate lines) and ground color (separate graphs) of the display for each cut, as a function of the hue difference between the figure and ground colors.

As described in the main article, pair-wise comparisons of cut combination 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 ($t(47) = 3.64$, $p < .002$).

Figure S2 shows analogous plots of harmony ratings for figure and ground cuts as a function of hue difference between the figure and ground colors. Similar to pair preference ratings, harmony ratings decreased monotonically as the hue difference between the figure and ground colors increased (Δ1-4, all cuts: $F(3,141) = 54.83$, $p < .001$; Δ0-4, different cuts: $F(4,188) = 111.70$, $p < .001$), but the reductions are more pronounced (Δ1-4, all cuts: $F(3,141) = 27.71$, $p < .001$; Δ0-4, different cuts: $F(4,188) = 43.90$, $p < .001$) (compare Figures S1 and S2).

There was a 3-way interaction among figure cut, ground cut, and hue difference (Δ1-4, all cuts: $F(27,1269) = 3.26$, $p < .001$; Δ0-4, different cuts: $F(24,1128) = 6.03$, $p < .001$). Relative to the other figural cuts, saturated figures are less harmonious with muted grounds of similar hues, light figures are more harmonious with light grounds of contrasting hues, light figures are more harmonious with muted grounds of similar hues, and muted figures are more harmonious with light grounds of similar hues. There was also an interaction between figure cut and hue difference (Δ1-4, all cuts: $F(9,423) = 4.19$, $p < .001$; Δ0-4, different cuts: $F(8,376) = 4.73$, $p < .001$) in which harmony ratings for pairs including dark figures decrease more rapidly as hue difference increased, relative to the other cuts. There was no such difference between ground cut and hue difference (Δ1-4, all cuts: $F(9,423) = 1.06$, $p > .05$; Δ0-4, different cuts: $F<1$).

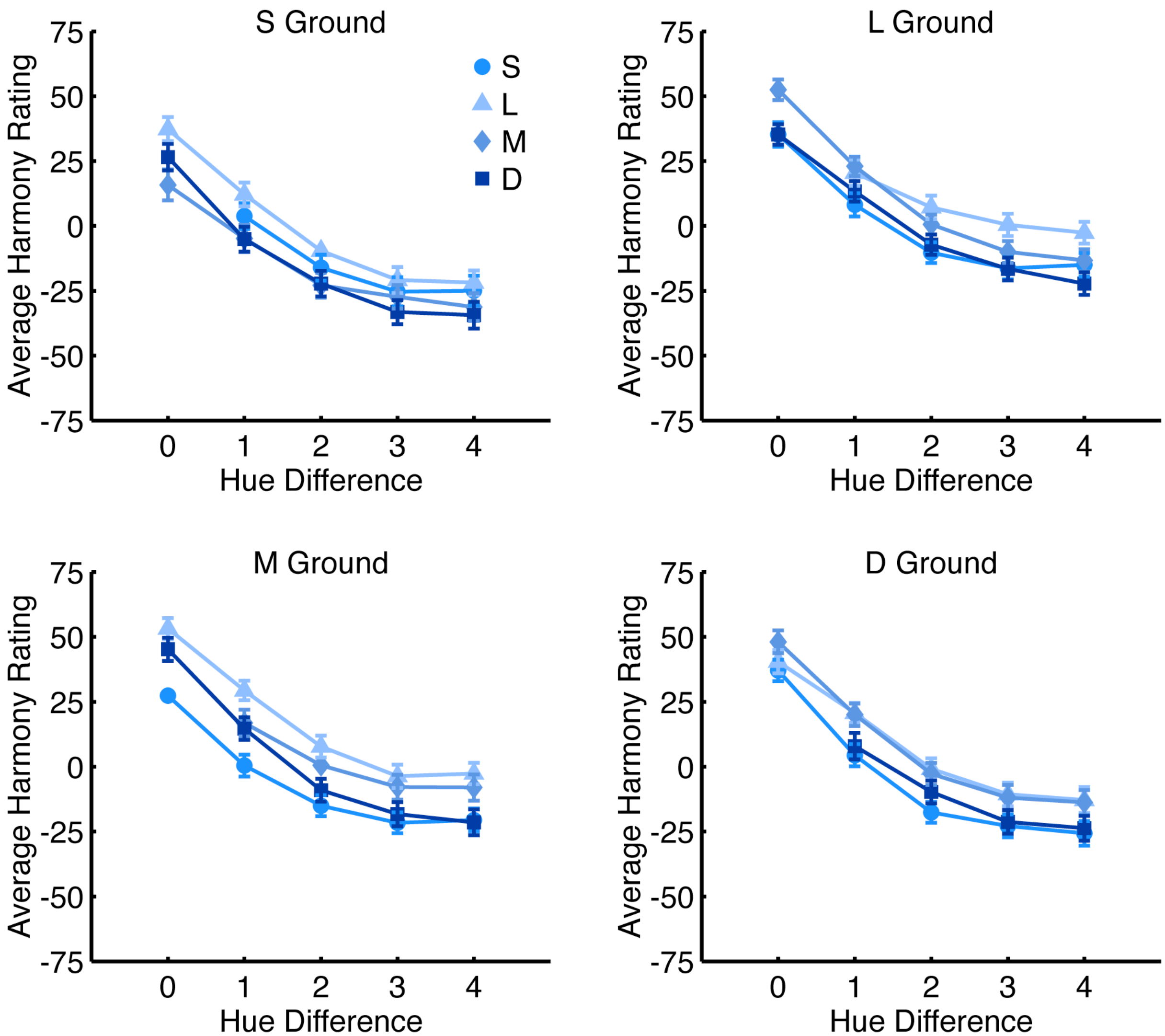

Figure S2. Harmony ratings for figural color (separate lines) and ground color (separate graphs) of the display for each cut, as a function of the hue difference between the figure and ground colors.

Pairwise comparisons between harmony ratings are shown in Figure S3. In summary, combinations that contained lighter and less saturated colors tended to be rated as more harmonious, in that L figures were judged most harmonious against all four ground cuts, and either D or S figures were judged least harmonious against all four ground cuts.

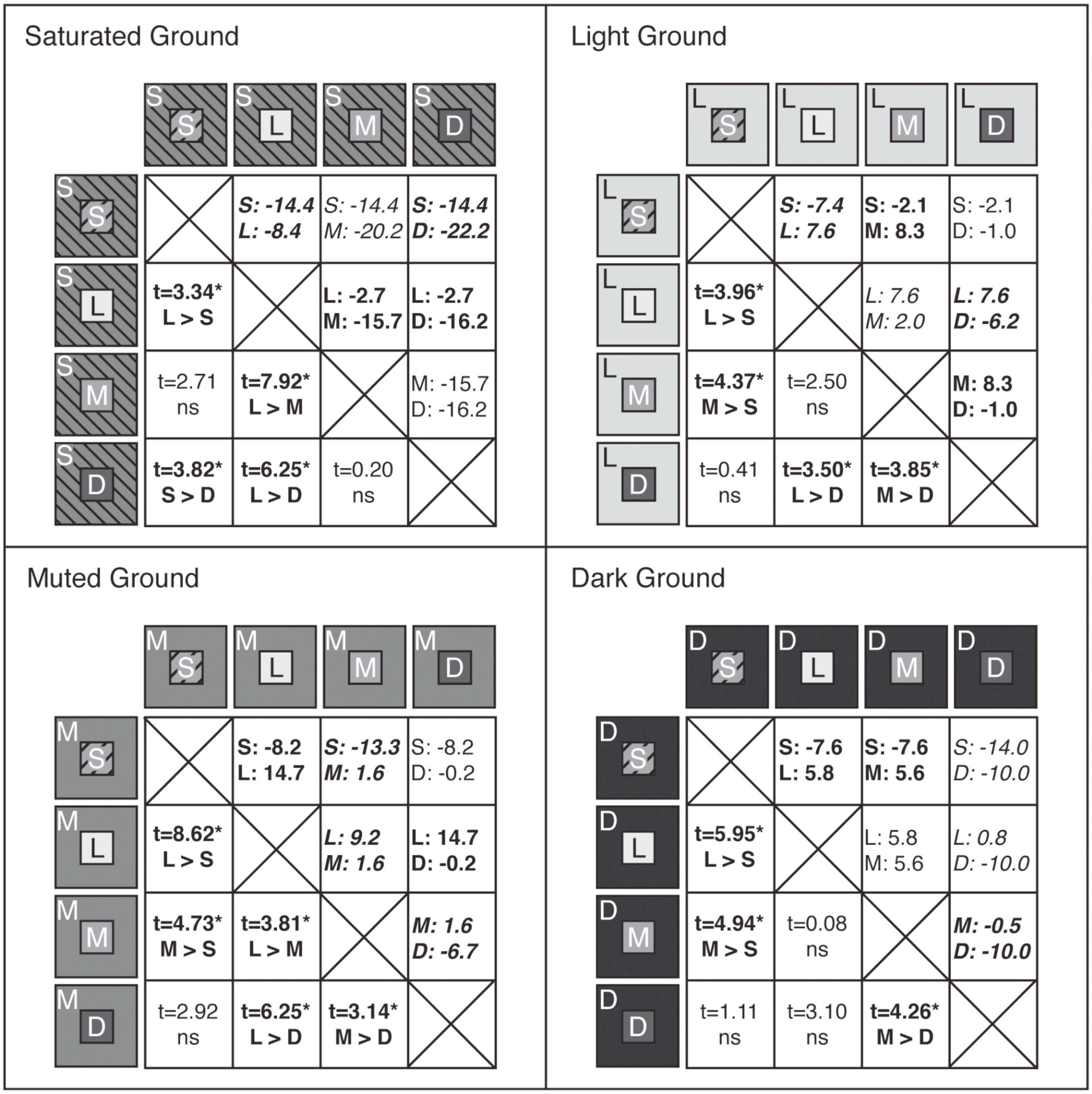

Figure S3. Comparisons between harmony ratings of color combinations with the same ground cut (separate quadrants) and different figure cuts. Icons adjacent to each row column represent the cuts of the figure-ground pairs that were judged. The lower triangle of each ground-cut matrix shows the results of t-tests (df = 47, *p ≤ .002, using the Bonferroni correction) and direction of the difference (e.g., "L > S" in the Saturated Ground quadrant indicates that light figures on saturated grounds were judged more harmonious than saturated figures on saturated grounds). The upper triangle of the matrix shows the means of the pairs that were compared. The means for comparisons with same-cut pairs (italicized text) include only hue-difference steps of 1-4 for both pairs. All other means include all hue-differences steps (0-4). Bold face text indicates differences were significant. (Note: In this diagram, the figure squares are colored slightly lighter than the ground to increase discriminability for same-cut pairs.)

Figure S4 shows pairwise similarity ratings for left and right region cuts as a function of hue difference between the two colors. Similar to pair preference and harmony ratings, similarity ratings decreased monotonically as the hue difference between the component colors increased (Δ1-4, all cuts: $F(3,141) = 311.08$, $p < .001$; Δ0-4, different cuts: $F(4,188) = 429.59$, $p < .001$), but with even more pronounced reductions than pair preference (Δ1-4, all cuts: $F(3,141) = 175.62$, $p < .001$; Δ0-4, different cuts: $F(4,188) = 199.90$, $p < .001$) and harmony ratings (Δ1-4, all cuts: 56.38, $p < .001$; Δ0-4, different cuts: $F(4,188) = 50.75$ $p < .001$).

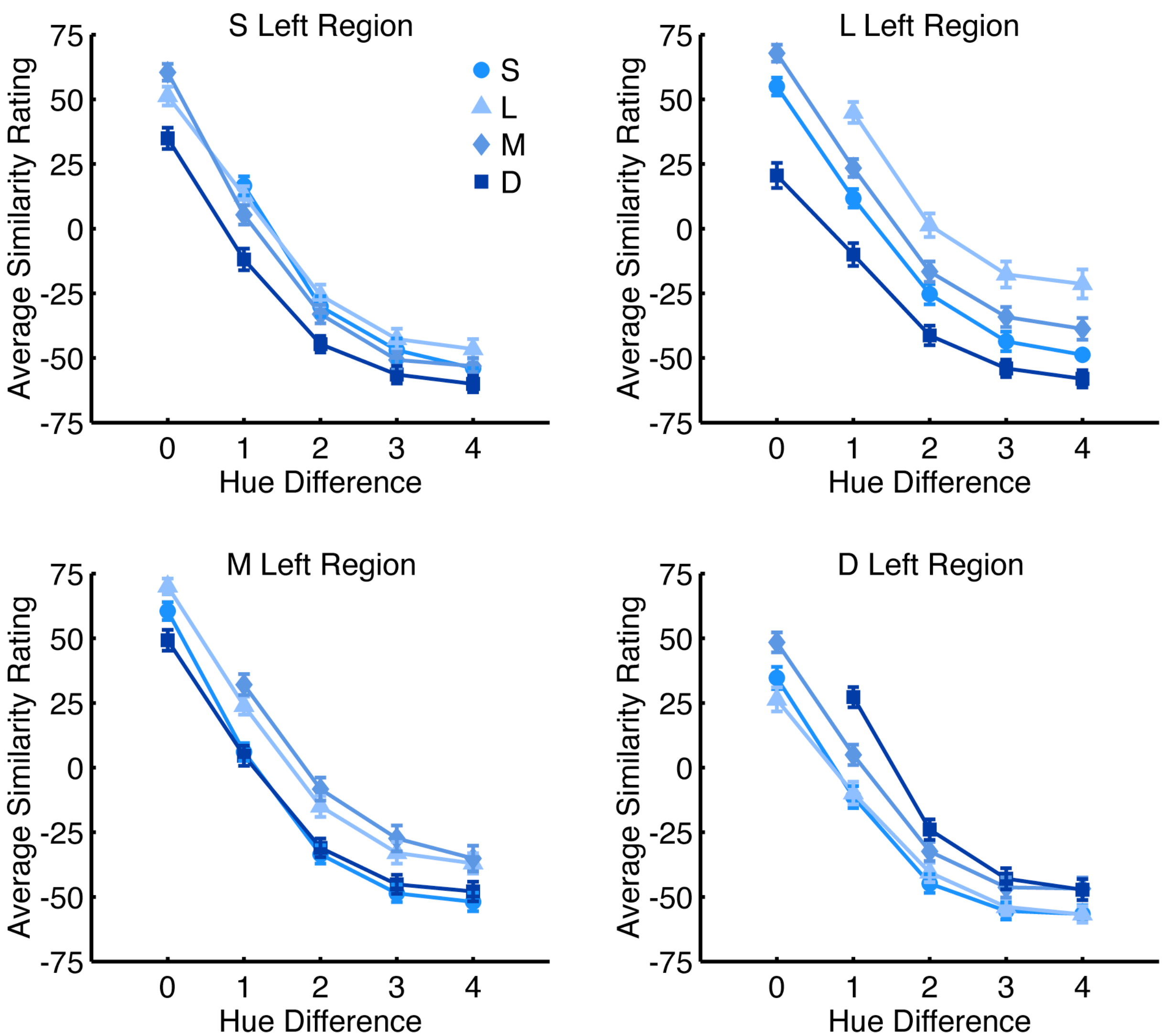

Figure S4. Similarity ratings for the left (separate graphs) and right (separate lines) region cuts, as a function of the hue difference between them.

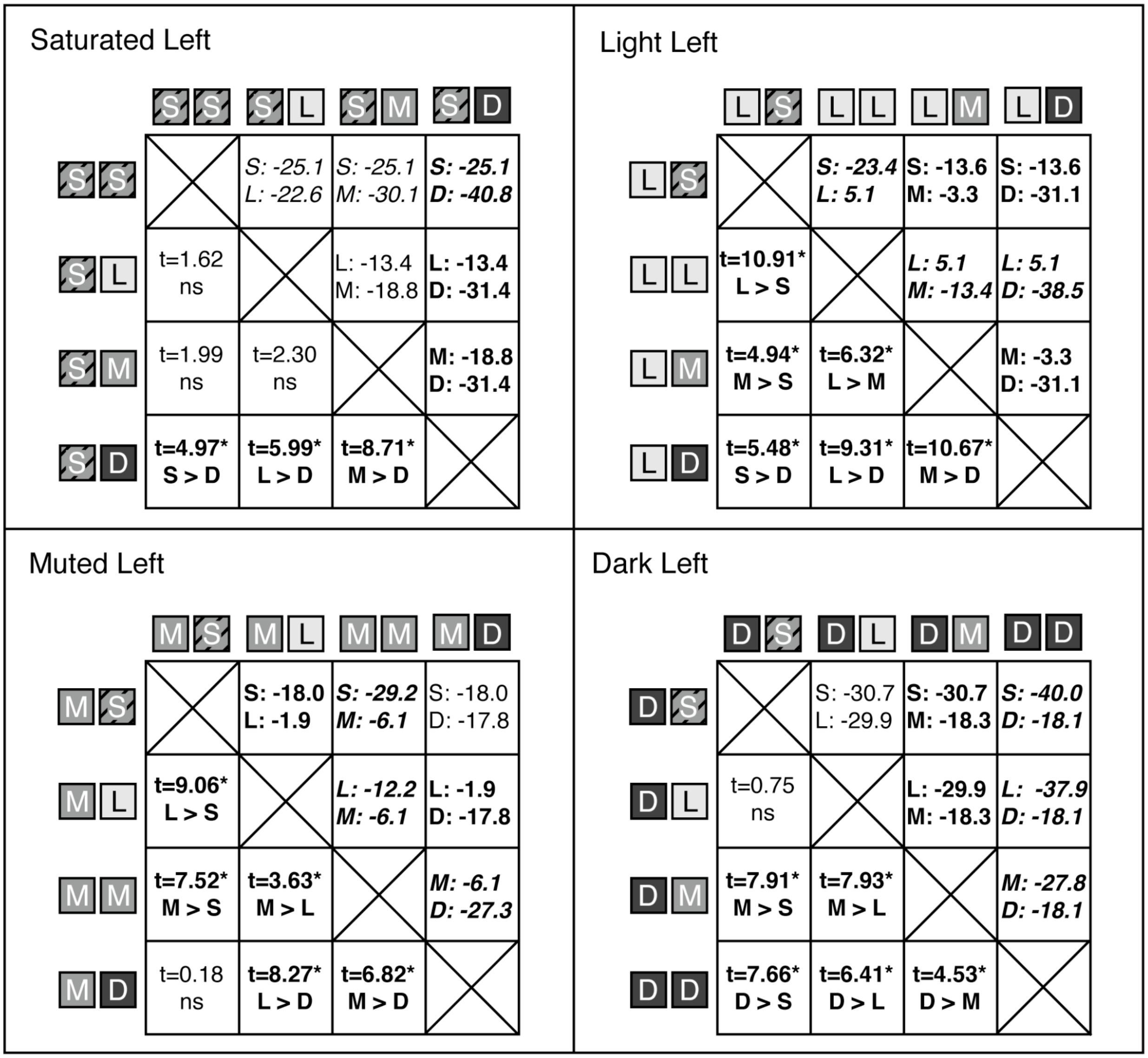

Figure S5. Comparisons between similarity ratings of color combinations with the same left region cut (separate quadrants in the figure) and different right region cuts. Icons adjacent each row and column represent the cuts of the pairs that were judged. The lower triangle of the matrix shows the results of t-tests (df = 47, *p ≤ .002 (using the Bonferroni correction) and direction of the difference (e.g., in the "Saturated Left" quadrant, "S > D" indicates that saturated colors (left) are more similar to saturated colors (right) than to dark colors (right)). The upper triangle of the matrix shows the means of the pairs that were compared. The means for comparisons with same-cut pairs (italicized text) include only hue-difference steps of 1-4 for both pairs. All other means include all hue-differences steps (0-4). Bold face text indicates differences were significant.

There was a 3-way interaction (Δ1-4, all cuts: $F(27,1269) = 14.30$, $p < .001$; Δ0-4, different cuts: $F(24,1128) = 12.36$, $p < .001$), between left region cut, right region cut, and the hue difference between the two. Not surprisingly, color pairs with similar lightness were judged to be most similar (see Figure S5 for specific comparisons), which was different from harmony ratings in which pairs that were generally lighter were more harmonious (compare Figures 7 and 11 in the main article). Similarity ratings decrease at different rates as hue difference increases, which is also evident in minor interactions between hue difference and left color cut (Δ1-4, all cuts: $F(9, 423) = 3.03$, $p < .001$); Δ0-4, different cuts: $F(12, 564) = 16.55$, $p < .001$, excluding same-cut pairs) and right color cut ($F(9, 423) = 2.1$, $p < .05$, $F(8, 376) = 12.73$, $p < .001$, excluding same-cut pairs).

Unlike pair preference, harmony and similarity ratings, figural color preference ratings did not decrease as hue difference between the figure and ground increased for all cut combinations (Figure S6). There was a 3-way interaction between figural cut, background cut, and hue difference (Δ1-4, all cuts: $F(27, 1269) = 8.68$, $p < .001$; Δ0-4, different cuts: $F(24,1128) = 10.36$, $p < .001$) as well as 2-way interactions between background cut and hue difference (Δ1-4, all cuts: $F(9, 423) = 8.30$, $p < .001$; Δ0-4, different cuts: $F(12,564) = 23.32$, $p < .001$) and figural cut and hue difference (Δ1-4, all cuts: $F(9,423) = 3.75$, $p < .001$; Δ0-4, different cuts: $F(8,376) = 7.41$, $p < .001$). For pairs with similar lightnesses – the four same-cut pairs (S-S, L-L, M-M and D-D) and the two S-M pairs – figural color preferences actually increased as hue difference increased (Δ1-4: $F(3,141) = 22.32$, $p < .001$). Pairs containing S figures on L backgrounds and L figures on S backgrounds also showed this pattern more weakly (Δ1-4: $F(3,141) = 7.20$ $p <.001$), even though L colors were lighter than S colors.

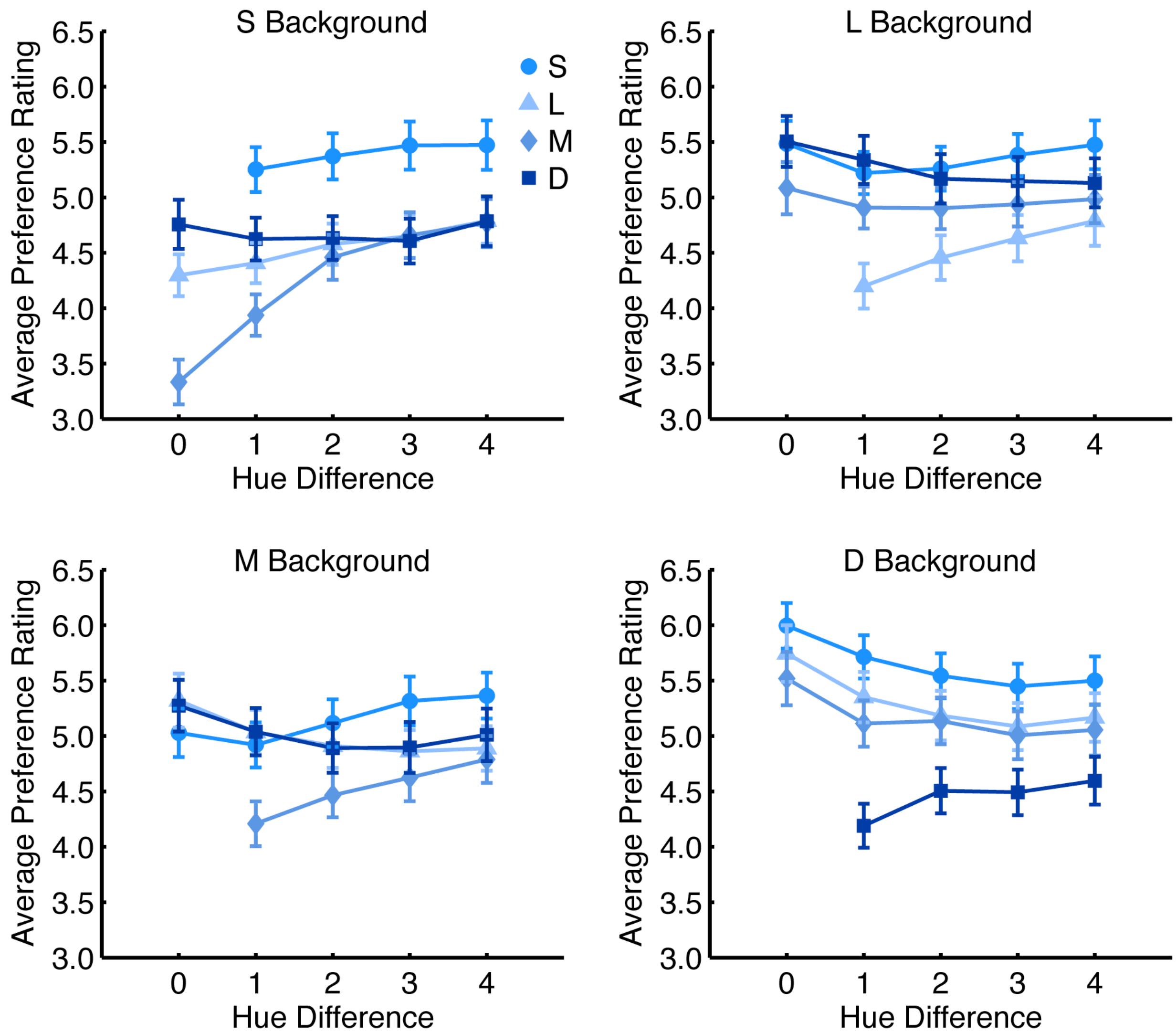

Figure S6. Preference ratings for figural colors (separate lines) on colored backgrounds (separate graphs), as a function of hue difference between the figural and background colors.

Further evidence for the importance of contrast in figural color preferences comes from the ordering of the cut effects (see Figure S7 for paired comparisons). On L grounds, the most contrastive D and S figures are most preferred and the least contrastive L figures are least preferred. Roughly the same is true for the D grounds, for which the most contrastive S and L figures are most preferred and the least contrastive D figures are least preferred. The M grounds are more contrastive with the S, L, and D figures than with the M figures, and the data follow this pattern as well (although the difference between the M and D figural colors was not significant after the Bonferroni correction was applied).

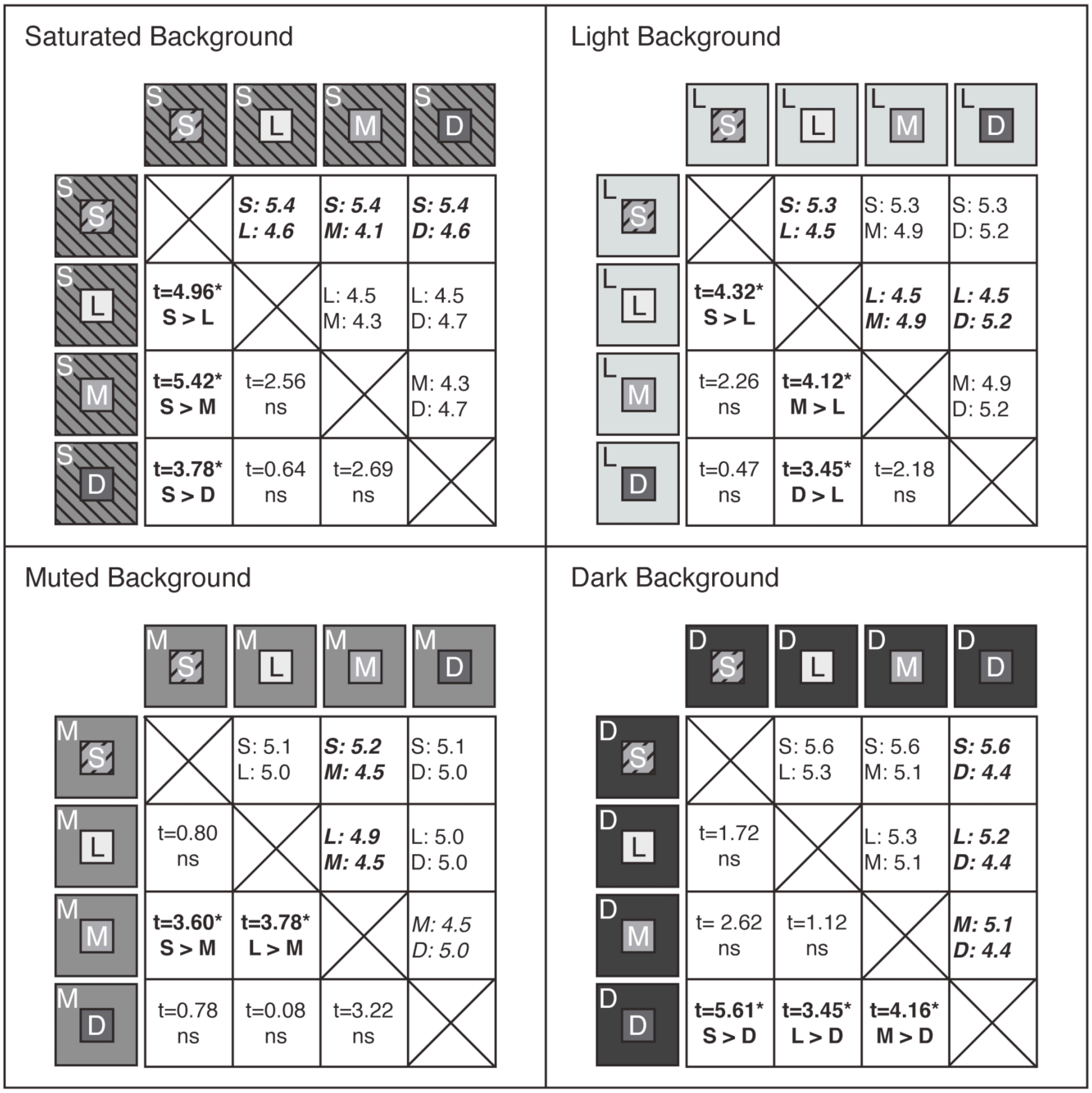

Figure S7. Comparisons between preference ratings for figural colors same background cut (separate quadrants) and different figure cuts. Icons adjacent to each row and column represent the cuts of the colors that were judged. The lower triangle of each background-cut matrix shows the results of t-tests (df = 47, *p ≤ .002 using the Bonferroni correction) and direction of the difference (e.g., "S > L" in the Saturated Ground quadrant indicates that saturated figures on saturated backgrounds were more preferable than light figures on saturated backgrounds). The upper triangle of the matrix shows the means of the pairs that were compared. The means for comparisons with same-cut pairs (italicized text) include only hue-difference steps of 1-4 for both pairs. All other means include all hue-differences steps (0-4). Bold face text indicates differences were significant.

The only results that appear to contradict this pattern are those for the S grounds, where the S figures are always most preferred even though it seems that these figures should have the lowest contrast with the ground. Notice, however, that highly saturated figures on highly saturated grounds will generally tend to be farther from each other in color space than the corresponding less saturated figures (L, M, and D figures) on highly saturated grounds.   Thus, it appears that most of the effects in figural color preferences can be attributed to some form of contrast, which generally enhances preference for the figural color.

Personality and Preference for Harmony

There was a high correlation between average pair preference ratings and average harmony ratings (r = +.79), but individuals participants varied widely in the degree to which they preferred harmonious color combinations (r ranged from -.03 to +.75). We tried to predict these individual differences in preference for harmony using the Big Five Inventory (or BFI), a 44 item personality inventory that measures five personality factors: Extraversion (talkative, assertive, active, energetic), Agreeableness (sympathetic, kind, appreciative, affectionate), Conscientiousness (organized, thorough, planful, efficient), Neuroticism (tense, anxious, nervous, moody), and Openness (wide interests, imaginative, intelligent, original) (John, Donahue, & Kentle, 1991; see John, Naumann, & Soto, 2008 for more detailed descriptions). Table S1 shows the correlations between participants' scores on each of the BFI factors and their degree of preference for harmony. Although none of the correlations even approached statistical significance, there are trends in which participants who showed a high preference for harmony were less extraverted and more conscientious.

Table S1. Correlations between preference-for-harmony measures and scores on the Big Five Inventory (BFI). Preference-for-harmony is measured for each participant by the correlation between his/her pair preference ratings and pair harmony ratings for all pairwise combinations of the 32 chromatic BCP colors.

BFI factor	Pearson's r	p-value
Extraversion	-0.16	0.27
Agreeableness	-0.01	0.93
Conscientiousness	0.13	0.38
Neuroticism	0.06	0.68
Openness	-0.09	0.55

When all five factors were tested in a linear regression, extraversion (2.7%, less extraverted, higher preference-for-harmony) and conscientiousness (an additional 2.3%, more conscientious, higher preference-for-harmony) were the factors that comprised the best fitting model (multiple-r = .22). Given that our sample size (n=48) is very small for personality research, further data must be collected to determine whether preference-for-harmony is related to these (or other) personality factors.

Formal color training and pair preferences

Figure S8 shows average pair preference ratings for participants with low (n = 7), moderate (n = 9), and advanced (n = 6) degrees of formal color training, respectively. The formal color training scale ranged from 1 (none at all) to 5 (very much), where 3 indicated training to the extent that traditional high school art offers. Participants in the low group scored a 1, in the moderate group scored a 3 and in the advanced group scored a 5, with those who scored a 2 or 4 being excluded to focus on the clearest cases.

To help understand the effects of such training, each group's data was analyzed separately using a regression model that included three factors: that group's average preference rating for the ground color, their average preference for the figural color, and their average harmony ratings (Figure S9). (These three factors were chosen to emphasize the differences in pair preferences between participants with different levels of formal color training, rather than to optimize the total percent of variance explained.)

Preference ratings in the moderate group decreased steadily as the hue similarity between the figure and ground color decreases (Figures S8 C-D). The results of the regression analysis showed that color harmony was indeed the most important factor for this group with only a small amount added by component color preference (see Formal Training group 3).

The preference function for participants with no formal color training appears to be more strongly driven by ground color preference, with pairs containing cool grounds being most preferred, and those containing warm grounds being least preferred (Figures S8 A-B). The same three-factor regression model showed that component color preference component color preferences are relatively more important for those with little or no formal color training than those with moderate color training (compare Formal Training groups 1 and 3 in Figure S9).

Interestingly, participants with advanced color training have generally flatter preference curves (Figure S5 E and F), indicating that they are less influenced by hue difference (and harmony) between the figure and ground color than the other two groups. As shown in Figure S9 (Formal Training group 5), only a moderate amount of variance in their preferences can be explained by their harmony ratings, with small amount added from component color preference. Perhaps the group with the most formal color training, who tended to be color professionals of various sorts (painters, designers, decorators, etc.) have had so much experience in working with color combinations that they have become bored with harmonious combinations and have come to appreciate contrastive combinations, in which hue differences are greater. Nevertheless, we see little evidence for Chevreul’s (1839) claim that highly contrastive colors are well liked even among this group of people highly trained in color theory.

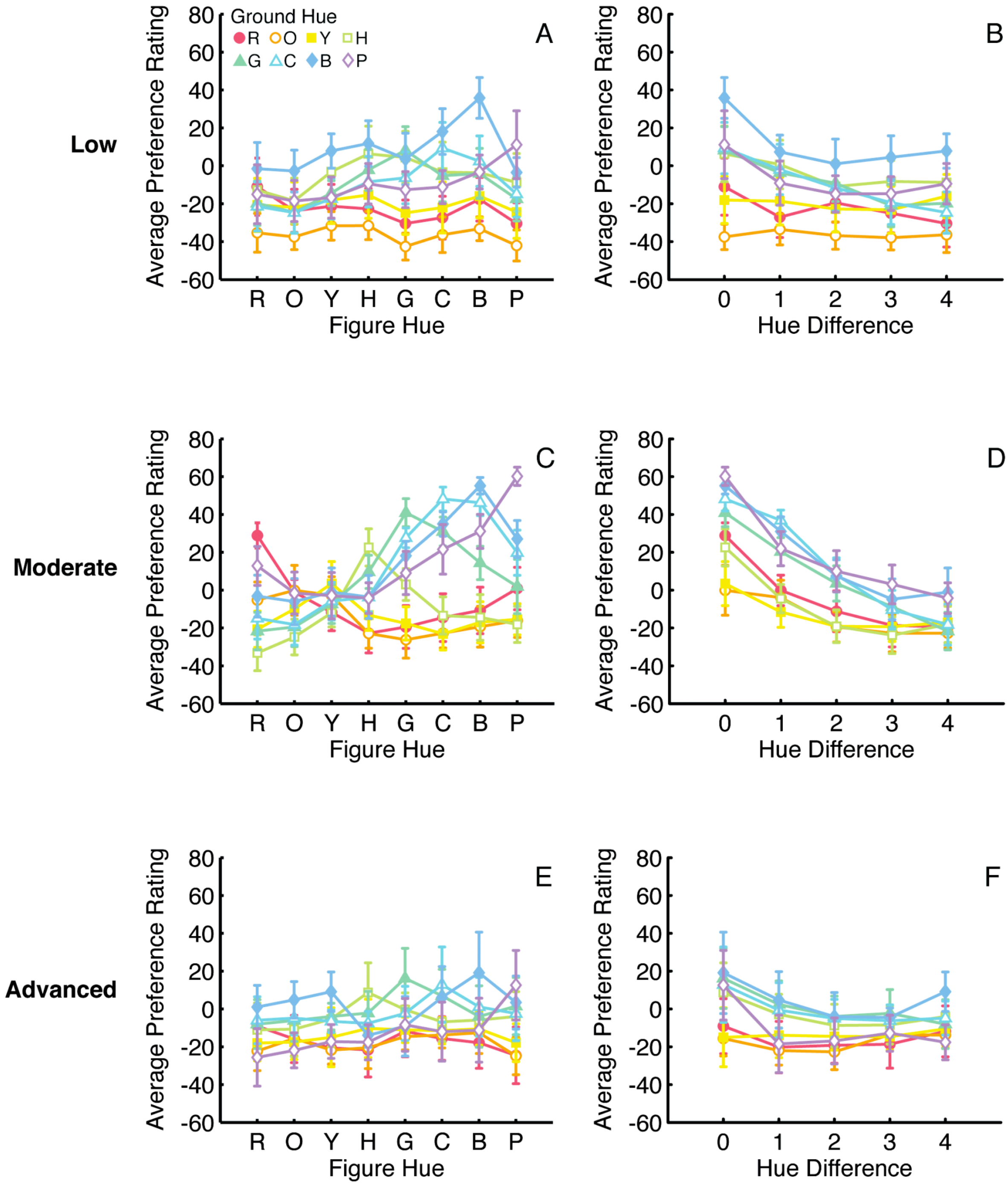

Figure S8. Pair preference for those with *low* (A-B), *moderate* (C-D) and *advanced* (E-F) formal color training, as a function of figural hue (x-axis) and ground hue (separate lines) (A, C, E) and as a function of the hue difference (in the present BCP design) between the figure and ground (B, D, F). Error bars represent the standard errors of the means (SEM).

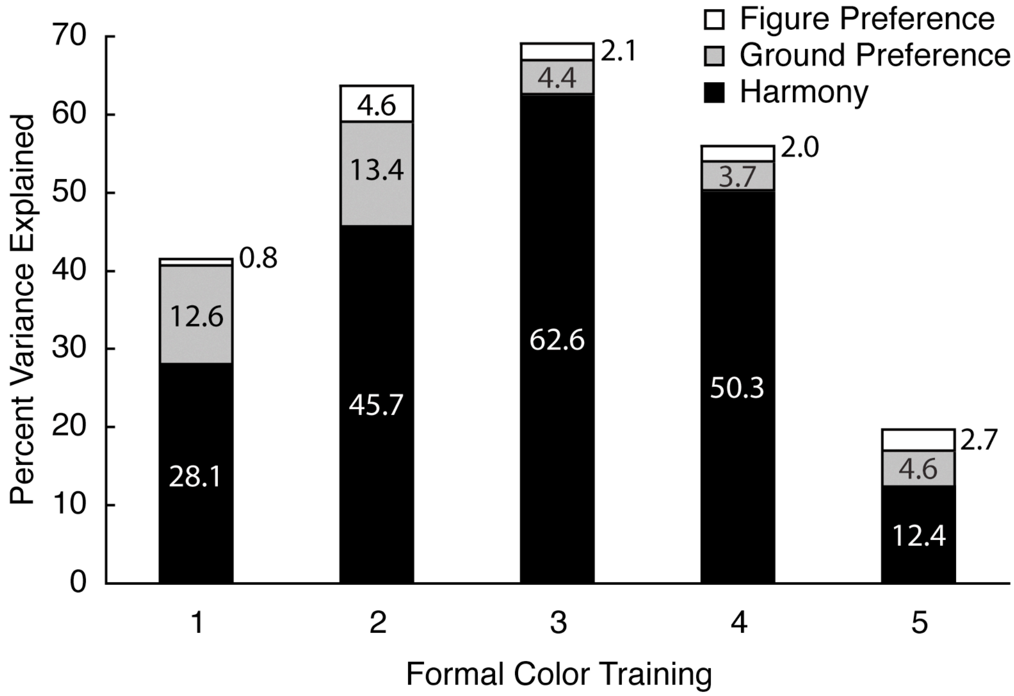

Figure S9. Bars show the percentages of variance explained by the harmony ratings (black stripe), ground color preference ratings (gray stripe) and figural color preference ratings (white stripe) for participants at each level of formal color training (1=none, 5=advanced). The order of the stripes represents the order in which each factor was entered into the regression model (bottom to top).

Preference, Harmony and Similarity Ratings as a Function of Hue Angle

Figures S10-S12 show the same hue functions, averaged over cut, for pair preference, harmony and similarity that are plotted in Figures 2, 6, and 10 of the main article, respectively, but replotted in terms of hue angle in CIELAB space. Figures S10A-12A show the same basic patterns as those in Figures 2A, 6A, and 10A: peaks where the two component hues are the same and decreasing roughly monotonically as hue difference increases. Figures S10B-S12B are slightly different from Figures 2B, 6B, and 10B because they show all eight color pairs for each ground hue (i.e., unfolded hue difference functions) rather than the average of clockwise and counter-clockwise hue differences (i.e., folded hue difference functions). In Figures S10B-S12B, positive hue differences are counter-clockwise and negative hue differences are clockwise around the color circle (see Figure 1B in the main article). The data are plotted so that hue differences were never greater than 180 deg in CIELAB space, which resulted in different numbers of data points to the left and right of the zero point for each curve.

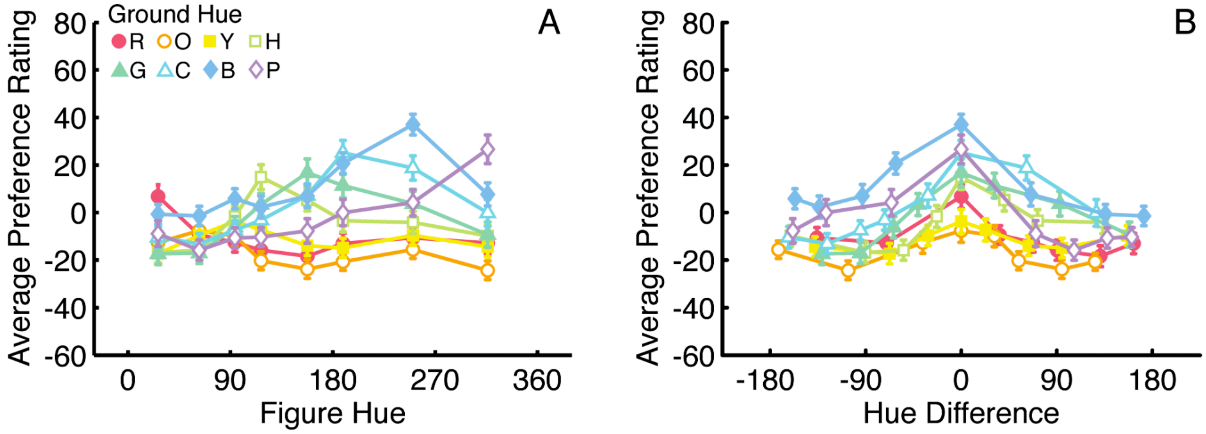

Figure S10. Preference ratings for pairs as a function of figural hue angle in CIELAB space (x-axis) and ground hue (separate lines) (A) and as a function of the hue difference between the figure and ground (+ hue difference indicates counter-clockwise and – hue difference indicates clockwise around the color circle in CIELAB space, see Figure 1B). (B). Error bars represent the standard errors of the means (SEM).

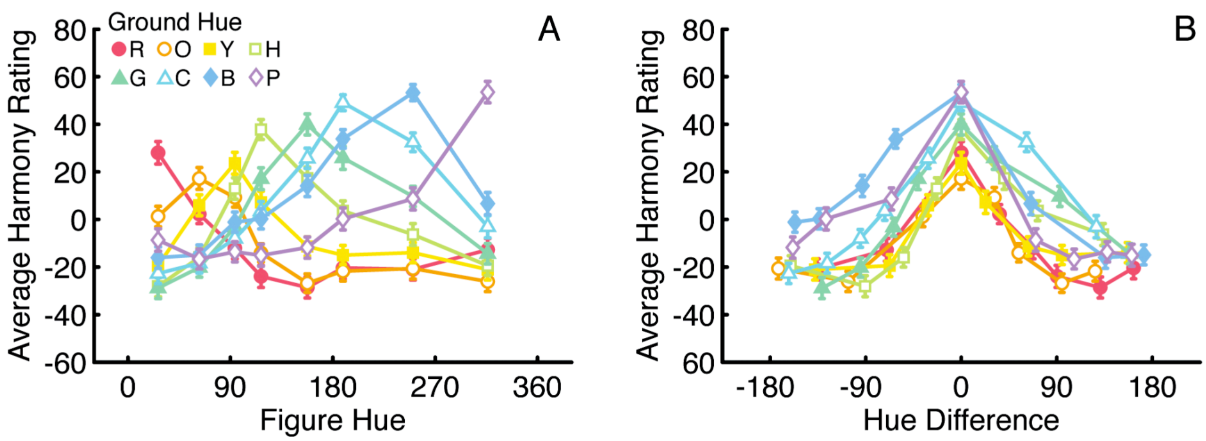

Figure S11. Harmony ratings for pairs as a function of figural hue angle (CIELAB) (x-axis) and ground hue (separate lines) (A) and as a function of the hue between the figure and ground (+ hue difference indicates counter-clockwise and – hue difference indicates clockwise around the color circle in CIELAB space, see Figure 1B) (B). Error bars represent the standard errors of the means (SEM).

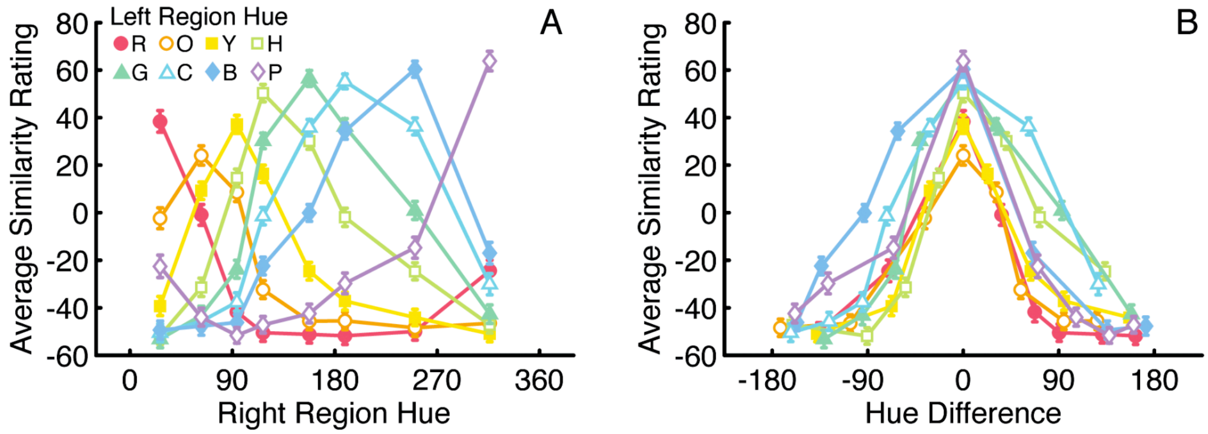

Figure S12. Similarity ratings for pairs as a function of right region hue angle (CIELAB) (x-axis) and left region hue (separate lines) (A) and as a function of the hue difference between the right and left regions (+ hue difference indicates counter-clockwise and – hue difference indicates clockwise around the color circle in CIELAB space, see Figure 1B) (B). Error bars represent the standard errors of the means (SEM).

Figure S13A shows the same figural color preference data as Figure 12A in the main article, but plotted as a function of figural color hue angle averaged over cut. Figure S13B shows the residual figural color preferences, (after accounting for figural color preference on a gray background and pair preference), as shown Figure 12B, but plotted as a function of figural color hue angle.

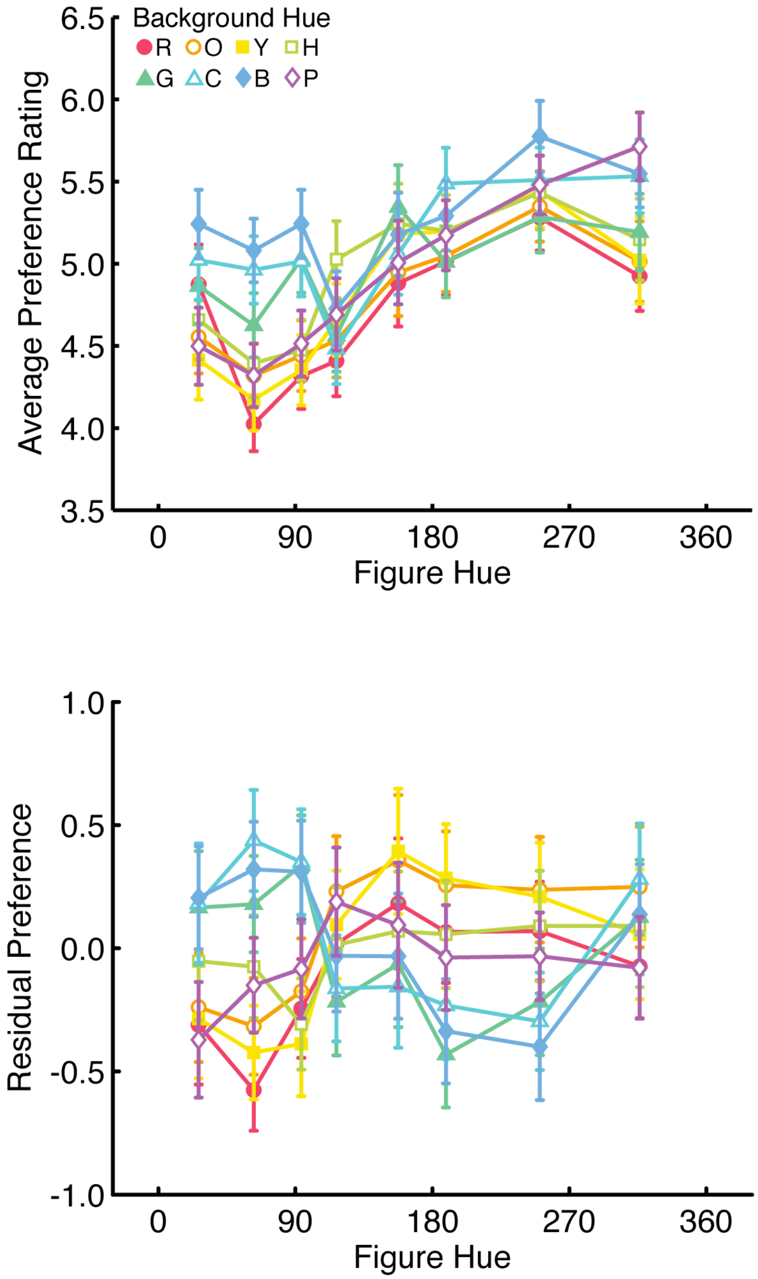

Figure S13. Preference ratings for each figural hue on each of the background hues as a function of figural hue angle (CIELAB) (A) and residual figural color preference after accounting for figural preferences when rated on a neutral gray background (Palmer & Schloss, in press) and pair preferences as a function of figural hue angle (B). Error bars represent the standard errors of the means (SEM).