Perception, 2011, volume 40, pages 635 – 648

doi:10.1068/p6655

Color relations increase the capacity of visual short-term memory

Thomas Sanocki, Noah Sulman
Department of Psychology, University of South Florida, Tampa, FL 33624, USA;
e-mail: Sanocki@usf.edu
Received 22 January 2010, in revised form 2 June 2011; published online 28 July 2011

Abstract. Do color relations such as similarity or harmony influence the ease with which colored patterns can be perceived and held in mind? We tested the influence of a relation supported in research on color harmony—similarity of hue—on the capacity of visual short-term memory (VSTM) for colors in patterns. Palettes of 4 similar-hue colors were rated as more pleasant (harmonious) than dissimilar-color palettes. The palettes were used in a VSTM color task. Patterns of 9 to 15 colored squares were presented, and accuracy of color change detection was measured. Memory performance was higher overall for similar-color palettes than for dissimilar-color palettes (experiments 1 and 3). Is this due to color similarity per se, or due to the harmony between colors in similar palettes? A final experiment provided strong support for the importance of color similarity as opposed to harmony. Overall, the advantages for color similarity, in terms of number of color squares held in memory (memory capacity) were 26% to 45% over dissimilar colors. The results indicate that color relations can have a strong impact on the capacity for perceiving and retaining color patterns.

1 Introduction

Is there a relationship between the appearance of objects and the efficiency with which they are processed in mind? There is strong evidence of such a relation in the domain of form perception where, in general, more elegant forms are easier to perceive, learn, and remember (Biederman 1987; Garner 1974; Palmer 1983). Color relations within patterns were of interest here. We presented patterns of colored squares, and used a short-term memory paradigm to ask: "Do color relations influence the ease with which the patterns are perceived and maintained in mind?".

The experiments were motivated by the hypothesis of color harmony. For nearly two centuries, there has been strong evidence of subjective preferences for certain types of color relations and, in particular, evidence of a preference for colors that are harmonious with each other (eg Chevreul 1839/1987; Granger 1955; Moon and Spencer 1944). Harmonious color palettes have been defined in a variety of ways, including tonal scales, complementary hues, and similar hues. However, the strongest evidence is for preferences based on similar hue (eg Chevreul 1839/1987; Chijiiwa 1987; Schloss and Palmer 2011; cf Arnheim 1974; Sloane 1989).(1) Our hypothesis was that patterns composed of similar hues would be easier to perceive and hold in visual short-term memory (VSTM).

Recently, however, evidence of a somewhat different relation between appearance and VSTM has been found (Lin and Luck 2009)—namely, that color similarity contributes to VSTM for individual colors. Color similarity may aid VSTM representation (Johnson et al 2009) and could explain an advantage for similar hues. In experiments 1 – 3 we collapse these competing color hypotheses, into the color-similarity hypothesis. We begin to distinguish between similarity and harmony in experiment 4.

---

(1) Color preferences are influenced by multiple factors, including physical properties such as stimulus size and composition, and general factors such as situation and task (eg Arnheim 1974; Ou and Luo 2006; Polzella and Montgomery 1993; Schloss and Palmer 2011; Sloane 1989). With color patches, which are of interest here, contemporary research converges on the conclusion that the most reliable color preferences based on harmonic relations are for colors that are similar in hue (Ou and Luo 2006; Poggesi et al 2009; Polzella and Montgomery 1993).

To test the effects of color similarity on VSTM, we designed 4-color palettes of high and low hue similarity. We confirmed that the similar-color palettes were viewed as more pleasant by observers, as predicted by the idea of color harmony. Then we used the palettes to generate patterns of colored squares, and measured observers' ability to perceive and hold the patterns in memory, with a standard measure of VSTM.

Short-term memory is a critical workspace for many mental processes, including comprehension, creativity, and visuo-spatial planning (Baddeley and Hitch 1974; Logie 2003; Potter 1993). VSTM is the medium for storing and manipulating visual and spatial information in mind (Logie 2003; Luck 2008). A central issue here is the maximum amount of information that can be held in VSTM—its capacity. There is extensive evidence that the capacity of VSTM is sharply limited, to no more than a few isolated objects (Vogel et al 2001; Zhang and Luck 2008).

However, the objects that VSTM holds can be quite complex (Alvarez and Cavanagh 2004; Phillips 1974; Sanocki et al 2001, 2010). VSTM can hold much of a typical real-world layout (Sanocki et al 2010), or a pattern involving as many as 17 independently varying location units (Phillips 1974; Sanocki et al 2001; Sanocki and Sulman 2008). The key to this ability appears to be relations between elements; memory is increased when the object's elements are more systematically related to each other spatially, by relations such as vertical or horizontal alignment (Sanocki et al 2010; Sanocki and Sulman 2008). This allows the elements to be grouped into larger and more mutually consistent hierarchical units (see also, eg, Jiang et al 2000; Miller 1956). Thus, observers may be able to hold several groups of elements in memory. This means that VSTM capacity for holding separate objects is limited, but the complexity of objects in VSTM is an open question.

In the present experiments we sought to further explore the impact of stimulus relations on VSTM capacity. Specifically, we asked if color relations also influence VSTM capacity. We measured VSTM capacity with the change-detection method. The stimulus patterns were defined by the color of each element or location [eg figure A1, see Appendix (in color online at http://dx.doi.org/10.1068/p6655)]; colors were randomly chosen from a palette. A memory pattern was presented briefly (200 ms), to be held in VSTM for 1 s. After the interval a test pattern appeared that was either identical to the memory pattern (*same*), or *different* in color shade at a single location. The color-similarity hypothesis is that accuracy of change detection would be higher with similar-color palettes than with dissimilar-color palettes.

The stimulus patterns consisted of 9, 12, or 15 square units that varied in color, allowing examination of VSTM as the demands on capacity increased. VSTM capacities, in terms of number of color units accurately held in memory, were estimated from the percentages correct using a simple high-threshold formula first applied by Pashler (1988) to change detection (see also, eg, Sanocki et al 2010; Vogel et al 2001). This model assumes that correct responses arise when the changed color unit is held in VSTM, or from correct guesses. The model produces a capacity estimate, number of color units (in this case) which is a straightforward measure of stimulus information held in memory. As noted, the units may be grouped into a smaller number of larger hierarchical structures. The model does not measure the number of groups held in memory.

The actual colors and color-changes were equated across palettes. The palettes consisted of 4 colors each in experiments 1 – 3, and dissimilar-color palettes were created by re-combining similar-palette colors (figure A2). To control changes across palette, the colors were organized as rows of similar (and same color-category) shades—row-pairs in figure A2. The critical changes (the differences between a memory pattern and its test pattern) were created by switching between the row-pair colors (eg the two greens from the top row in figure A2). This meant that the same color changes were used in the similar and dissimilar conditions. For example, the interchange of top-row green shades would occur equally often with the similar-color palette (top palette in figure A2) or with a dissimilar-color palette (4th in figure A2).

2 Experiment 1

The two similar-color palettes consisted of shades of red and of shades of green. The dissimilar-color palettes combined rows of reds and greens (figure A2). Red and green are distant values on the color wheel and opponent codes in color vision. Red – green opponency may have evolved in part because of advantages in discriminating fruits from vegetation (eg Mollon 2000). The expected differences in harmony were validated by independent ratings. In this experiment, palettes were presented in separate blocks for the VSTM task.

2.1 *Method*

2.1.1 *Participants.* Students from introductory psychology courses at the University of South Florida participated in exchange for course credit. Data from twenty-one participants (seventeen females) were analyzed in the memory experiment. One dataset was omitted because the participant had low VSTM accuracy ($< 60\%$), possibly because of color vision deficiency (which was not measured separately from VSTM performance). A separate group of twelve participants (nine females) rated the palettes for color harmony; they reported normal color vision.

2.1.2 *Stimulus design and procedure.* Stimulus patterns consisted of 3 rows of 3, 4, or 5 square units (locations), centered on the screen. For a given trial, each stimulus unit in the memory pattern was randomly assigned a color from the relevant palette. Test patterns were identical on *same* trials. For *different* trials, a single unit was randomly selected and its color replaced by the other member of the row-pair (rows in figure A2, as explained above). At the start of each trial, a ready signal (cross) was presented for 500 ms, followed by the memory pattern (200 ms), a 1 s blank interval, and then the test pattern until the response (*same* or *different*, using ‘1’ or ‘2’ on the numerical keypad, respectively). Visual angle of the stimulus patterns ranged from 17.2 deg × 17.2 deg (smaller display size), to 28.5 deg × 17.2 deg (larger display size, horizontal by vertical), at the viewing distance of approximately 60 cm. Because of the complex displays and brief presentation duration for the first array, observers were encouraged to encode the array as a whole pattern. This is consistent with the idea that complex patterns can be quickly encoded and held in VSTM (eg Phillips 1974; Sanocki et al 2010).

Trials were organized in 8 blocks of 24. Display size and response were crossed within blocks. Block order rotated through the 4 color conditions (2 similarities × 2 palettes) in an order counterbalanced across subjects. Thus, there was a total of 192 test trials and 8 observations in each factorial condition. The relevant palette was presented for inspection during the rest period before the start of a block. Two practice blocks (one similar-color, one dissimilar-color) preceded test; for the first block, durations of the memory patterns were increased to 500 ms, to make the stimulus sequence easier for the observers to learn. The experiment was controlled by a Macintosh G4 using custom software (Realbasic), with a 17-inch CRT Macintosh monitor.

2.1.3 *Harmony ratings.* The separate group of observers was instructed to rate ‘color harmony’, which was defined as ‘a pleasing relation between colors’, on a 7-point scale. The palettes were presented as 4 elements arrayed as in figure A2. The observers participated in a single group session, and saw all palettes; each palette was presented once for a practice rating and then once again for a single test rating. Mean test ratings are shown in figure A2. Ratings were reliably higher for the similar-color palettes ($M = 5.9$) than for the dissimilar-color palettes ($M = 2.4$; within-observer standard error $= 0.26$) ($F_{1,11} = 171.06$, $p < 0.001$). Differences between palettes nested within harmony were also reliable ($F_{2,22} = 35.84$, $p < 0.001$, see figure A2).

2.2 *Results*

Figure 1 shows accuracy as a function of display size and response in the two main color conditions. As can be seen, overall performance was higher for the high-similarity palettes (gray lines in figure) than for low-similarity (darker lines), for both responses. Analysis of variance (ANOVA) confirmed the overall advantage for similar-colors over dissimilar-colors ($F_{1,20} = 11.42$, $p = 0.003$, $\eta_p^2 = 0.36$). Responses were more accurate on *same* trials ($F_{1,20} = 63.21$, $p < 0.001$, $\eta_p^2 = 0.76$), and an interaction between color similarity and response approached reliability ($F_{1,20} = 3.71$, $p = 0.07$, $\eta_p^2 = 0.16$). As can be seen in figure 1, similar-color palettes increased accuracy of *different* responses (6.3% advantage, $F_{1,20} = 63.21$, $p < 0.01$) but not *same* responses (0.7% advantage, $F_{1,20} < 1$). Color similarity may aid detection of differences, as explained below and as predicted by a model of similarity effects in VSTM (Johnson et al 2009). There were also effects of display size ($F_{2,40} = 34.10$, $p < 0.001$, $\eta_p^2 = 0.63$) and an interaction of display size and response ($F_{2,40} = 7.09$, $p = 0.002$, $\eta_p^2 = 0.26$).

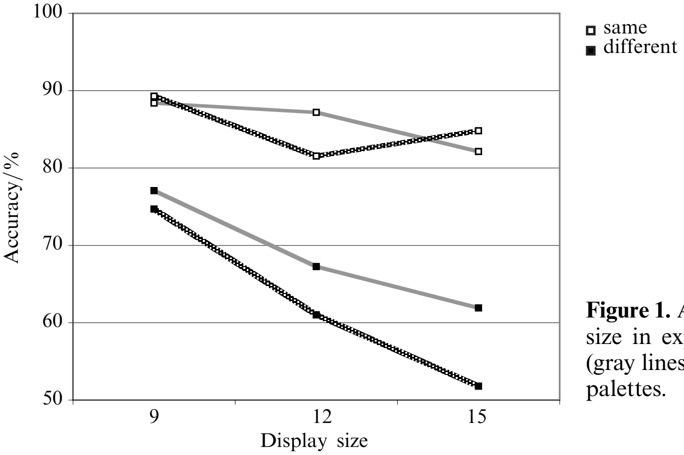

Figure 1. Accuracy as a function of display size in experiment 1, for high-similarity (gray lines) and low-similarity (dark lines) palettes.

Because each palette had unique color relations, the factor palette was nested within color similarity. Palette was involved in no reliable effects; overall performance was similar for the 2 similar-color palettes (greens: 77.9%; reds: 76.8%), and for the 2 dissimilar-color palettes (reds/pale greens: 73.7%; greens/pale reds: 74.0%).

There was a tendency for advantage for color similarity to increase with display size (figure 1), although the interaction of similarity and display size was not reliable ($F_{2,40} = 1.67$, $p > 0.20$, $\eta_p^2 = 0.08$). Such an effect would be potentially interesting because it would suggest that color similarity increasingly helps performance as capacity limits are reached.

Pashler's (1988) k is an estimate of the number of color units accurately held in memory,(2) and is most valid when calculated from the largest display size. Estimated capacity was 7.9 stimulus units for similar-color conditions and 6.3 units for dissimilar-color conditions. The difference in k values was reliable ($t_{20} = 2.23$, $p = 0.04$, two-tailed, SED $= 0.73$). The gain in stimulus capacity with similar-color relations was 25.8%.

2.3 *Discussion*

The results provide tentative support for the prediction that VSTM capacity, in terms of number of stimulus units, will be greater with similar-color palettes. The 26% gain in capacity with similar-color is sizable and a potentially significant gain in performance in the real-world.

(2) Pashler's k estimates this from the hit and false alarm rates, under the assumption that correct *different* responses come from either (a) the changed item being held in VSTM, or (b) guesses (*different*) on the remaining portion of change trials (see Vogel et al 2001; Pashler 1988).

However, there are at least several types of alternative explanations of VSTM performance. First, interactions occur between adjacent, simultaneously presented color regions. Although the critical color discriminations were equated between color conditions, the discriminations might be aided by having similar colors in the immediate context, relative to having dissimilar colors. In other research paradigms, color discrimination is more accurate in contexts of similar color (Smith and Pokorny 2003). However, the present paradigm differs from discrimination paradigms in a number of ways. We tested for simultaneous color discrimination effects in experiment 2, by measuring the critical color discriminations in similar-color and dissimilar-color contexts. To anticipate, we found that simultaneous color discrimination was somewhat easier in the dissimilar-color contexts.

A second concern is homogeneity of the colors within displays. Could homogeneity be the determinant of higher performance? Note that homogeneity of hue is consistent with our definition of color similarity. On the other hand, homogeneity of luminance can be distinguished from hue similarity. Luminance values varied within palettes but the variation was similar between color conditions: the standard deviation of the 4 luminance values in each palette (L values in figure A2) averaged 18.7 for the similar-color palettes and 18.9 for the dissimilar-color palettes, on a 100 unit scale. Thus, homogeneity of luminance does not explain the advantage for similar-color palettes in experiment 1.

A third factor to consider is number of color categories within stimulus patterns. Whereas the similar-color palettes each involved one primary category, the mixed palettes included two color categories in experiment 1, green and red. Memory capacity could have been higher with similar-color patterns because there were fewer color categories in the patterns. In experiment 3, we redesigned the palettes to all consist of two color categories. Also, to minimize the use of palette-specific strategies, we intermixed palettes during the experiment, in contrast to the blocked presentation in experiment 1.

The color similarity effect in experiment 1 tended to be greater on *different* trials than on *same* trials. This result is predicted by a neural-network model of similarity effects on VSTM proposed by Johnson et al (2009). In the model, color similarity enhances the detection of color changes, as will now be explained. The model represents the initial array as a layer separate from the second, test, array and there is competition between layers (inhibitory input from the initial layer to the second layer). Within the initial layer, similar colors inhibit each other more than dissimilar colors. This lowers activation in the initial layer, and as a result reduces inhibitory input to the second layer. This enhances detection of changes present in the second array.

3 Experiment 2

The purpose of this experiment was to provide a test of the idea that simultaneous color discrimination may be aided by high-similarity color contexts. Our aim was to measure simultaneous color discrimination between the colors that changed (the row-pairs) in experiment 1, while varying the surrounding color context. One stimulus pattern was presented on each trial, and it consisted of a critical middle row surrounded by context rows above and below (figure A3). Observers indicated whether the two units of the middle row were the same exact color or two colors. When the colors differed, they were the two members of a row-pair. The manipulation was the irrelevant top and bottom rows of the pattern; they were filled (randomly) from the other two colors of the similar-color or dissimilar-color palette. We presented the patterns briefly to make the discriminations difficult, and measured accuracy of discrimination.

3.1 *Method*

The 4 palettes of experiment 1 (figure A2) were used equally often and in a random order. Palette was selected (randomly) before each trial. Then a color was randomly selected from the palette to be used in the critical middle row, along with its row-mate (from figure A2) on *different* trials. The remaining 4 stimulus units (top and bottom rows) were each filled by randomly selecting from the 2 remaining colors in the palette. Trials began with a 500 ms fixation cross, followed by the stimulus pattern centered in the same position. There was no mask. There were 8 blocks of 16 test trials each, preceded by 2 practice blocks. *Same* and *different* trials were equally likely. The first practice block had long stimulus durations (500 ms) to help observers learn the task; the second practice block had a stimulus duration of 150 ms. Thereafter, the duration was adjusted in a staircase method to produce accuracy near 75%. The average ending duration was 35 ms (SD = 23 ms). The dependent variable was percentage of correct discriminations at the brief durations. A total of twenty-eight new students (twenty-six female) participated.

3.2 *Results and discussion*

The mean percentage correct was 75.2% when the color discriminations were in a similar-color context and 78.6% in the dissimilar-color context ($t_{27} = 2.60$, $p = 0.02$, SE = 1.29%). Thus, the similar-color surround hurt simultaneous color discriminations. This result is inconsistent with an explanation of color-similarity effects on VSTM that emphasizes simultaneous color discriminations. Simultaneous discriminations between the changing colors in the VSTM experiment were not aided by the similar-color contexts. Explanations involving the perception and retention of color in VSTM are supported.

The result contrasts with the finding that color discrimination can be higher when the background colors are more similar to the target colors (Smith and Pokorny 2003). There may be a number of reasons for this difference, following from the many differences in methods. One potentially important set of factors includes the considerable change here in colors from trial to trial (each unit in the array changes), and the need to attend to the critical units separately from (ignoring) the other changing, irrelevant units. Color differences between the context units and critical units may enhance this selective attention process. In the main VSTM experiments, there is also a need to separate a changing element from other elements that change from trial to trial. Therefore, in the present context, simultaneous color discrimination does not seem to be a promising explanation of the accuracy differences in VSTM.

4 Experiment 3

In experiment 3 the influence of similarity on VSTM was measured again. In this experiment every palette had two color categories. To begin generalizing findings across colors, we used a new set of colors. The similar-color palettes consisted of the adjacent color categories blue – purple and yellow – green, while dissimilar-color palettes consisted of opposing pairings of these colors (figure A4). Example 9-element patterns are shown in figure A5. Also, to minimize possible influences of strategies when the colors are blocked (experiment 1), we varied the palette from trial to trial.

4.1 *Method*

There were three procedural changes from experiment 1. First, palette was randomized across trials of the experiment. Second, we added a constraint to the routine for selecting colors for stimulus patterns, to increase the mixing of the colors: each of the 4 colors in a palette was randomly assigned to an approximately equal number of locations. Because 2 display sizes were not divisible by 4, color frequencies differed by 1 within individual patterns but were equal on average.

The third change was that the harmony ratings were done by the memory participants, after the test trials were complete, on a 100-point scale. The palettes were presented for ratings as 12 element patterns (3 units of each color), once for practice and then once for test. Mean harmony ratings are shown for each palette in figure A4.

Ratings were reliably higher for the similar-color palettes ($M = 76.3$) than for dissimilar-color palettes ($M = 39.4$, within-observer standard error = 3.49, $F_{1,11} = 171.06, p < 0.001$). Also reliable were differences between palettes nested within harmony ($F_{2,22} = 35.84$, $p < 0.001$, see figure A4). A total of thirty-eight new participants (thirty-two females) provided data in this experiment (two others had low accuracy).

4.2 *Results and discussion*

Figure 2 shows accuracy as a function of display size and response in the two main color conditions. As can be seen, overall performance was higher for the high-similarity palettes than for low-similarity for both responses. ANOVA confirmed the overall advantage for similar-colors ($F_{1,37} = 14.05$, $p < 0.001$, $\eta_p^2 = 0.28$). The interaction of color similarity and response approached reliability ($F_{1,37} = 3.63$, $p = 0.064$, $\eta_p^2 = 0.09$). As can be seen in figure 2, similar-color palettes increased accuracy of *different* responses (5.9% advantage, $F_{1,37} = 35.06$, $p < 0.001$) but not *same* responses (1.5% advantage, $F_{1,37} = 2.19$, $p > 0.3$). This is consistent with predictions of the Johnson et al (2009) model of color VSTM. There were main effects of display size ($F_{2,74} = 37.89$, $p < 0.001$, $\eta_p^2 = 0.51$), and response ($F_{1,37} = 160.91$, $p < 0.001$, $\eta_p^2 = 0.81$), and an interaction of display size and response ($F_{2,74} = 3.76$, $p = 0.03$, $\eta_p^2 = 0.09$).

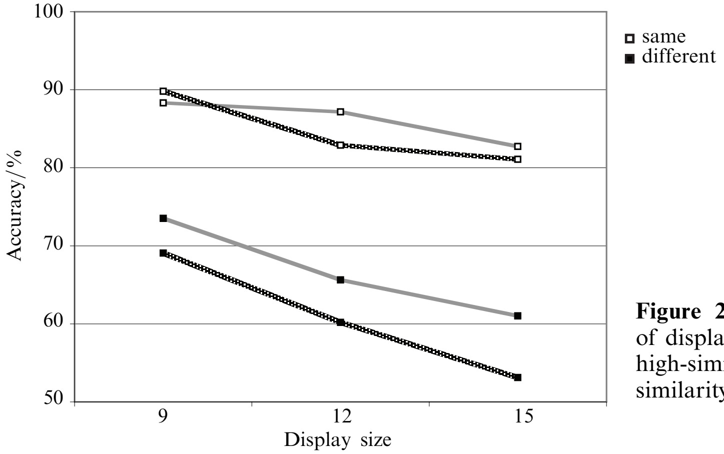

Figure 2. Accuracy as a function of display size in experiment 3, for high-similarity (gray lines) and low-similarity (dark lines) palettes.

In this experiment, there was also a main effect of palette nested in color similarity condition ($F_{2,74} = 10.32$, $p < 0.001$, $\eta_p^2 = 0.22$). Accuracy was similarly high for the two similar-color palettes (green – yellow, 75.7%, blue – violet, 77.1%; $t_{37} < 1$), but different for the two dissimilar-color palettes (violet – green, 69.6%, blue – yellow, 75.8%; $t_{37} = 4.54$, $p < 0.001$). The reason for the higher performance with the blue – yellow palette remains unclear. Note that ratings of harmony were higher for the violet – green palette than for the blue – yellow palette. None of the remaining effects involving palette was reliable.

There was again a tendency for the color similarity effect to increase with display size (figure 2). However, the interaction of color similarity and display size was not reliable ($p > 0.10$).

The capacity estimate was 7.9 stimulus units for similar-color palettes and 6.1 units for dissimilar-color palettes ($t_{37} = 3.25$, $p = 0.002$, two tailed, SED = 0.59). The increase in stimulus capacity with color similarity was 29.3%.

Is there a relation between ratings of palette harmony and memory performance at the individual observer level? Individual observer ratings were not very reliable because we collected only one rating per palette. Nevertheless, the relation is of some interest. Therefore, we correlated each observer's palette rating with his/her overall memory performance with that palette. The correlations averaged 0.18 across observers, suggesting a weak relation. Further research at the individual observer level is warranted; indeed, the relation between harmony and memory at the individual observer level is a critical issue for future research.

5 Experiment 4

The results so far indicate that similar hues are held in mind more efficiently than dissimilar hues. However, is the critical factor hue similarity, or is there something to the notion of color harmony? Can harmony be distinguished from similarity?

Our approach to the question begins with the idea that, if color harmony is important, then there should be a benefit of having additional harmonious colors in the display. For example, compared to 2-color displays, 4-color displays should provide increased color richness and more opportunity for harmonizing. An analogy to singing is relevant here—the sound of four voices provides a richer chorus than the sound of two voices. This idea contrasts with a prediction that follows necessarily from similarity: 2 colors can be more similar to each other than 4 colors. Specifically, two similar colors (A, A') will be more similar to each other than a combination of those two colors and two additional colors (B, B') from outside the $A-A'$ range. This holds for the palettes in figure A6, where 4-color palettes are compared to 2-color palettes of similar colors. The claim is that there is more similarity within each 2-color palette than within the 4-color palette.

To test these predictions, we compared VSTM performance between palettes of 4 harmonious colors, and palettes of 2 similar (and harmonious) colors. If harmony is more important than similarity, then additional harmonious colors should produce higher performance. On the other hand, if similarity is critical, then performance should be higher with 2-color palettes, because the overall inter-color similarity is necessarily higher than with 4-color palettes.

We tested these predictions with two different color-schemes, and two groups of observers. The 4-color palettes were two palettes highly rated for harmony in the previous experiments. Each color scheme consisted of (each group of observers saw) one 4-color palette and two 2-color palettes derived from the 4 colors (figure A6). Each group received the 4-color palette on half of the trials and each 2-color palette on one-quarter of the trials (in a randomized order).

5.1 *Method*

The method for the experiment was similar to that in experiment 3, with the addition of a second group of observers receiving a different color scheme. One color scheme was derived from the red palette from experiment 1 (figure A6, top), which was highly rated for harmony in that experiment. The other scheme was based on the yellow/green palette from experiment 2 (figure A6, bottom), chosen because its colors complemented the red palette and because it was highly rated for harmony in that experiment. Seventeen observers (thirteen females) received the red scheme and seventeen (thirteen females) received the yellow–green scheme. Harmony ratings were not collected because we were concerned that comparisons between 4-color and 2-color palettes would be confusing.

5.2 *Results*

Figure 3 shows the accuracy functions for the two different palette sizes. As can be seen, accuracy was higher for the 2-color palettes than for 4-color palettes for both responses. An overall ANOVA confirmed the large advantage for the more similar, 2-color palettes ($F_{1,32} = 335.04$, $p < 0.001$, $\eta_p^2 = 0.91$). There was also an effect of response ($F_{1,32} = 49.10$, $p < 0.001$, $\eta_p^2 = 0.60$), and a large interaction of palette size and response ($F_{1,32} = 190.13$, $p < 0.001$, $\eta_p^2 = 0.86$). The advantage for 2- over 4-color palettes was large for *different* patterns (29.7%; $F_{1,32} = 882.36$, $p < 0.001$) and small, but still reliable, for *same* patterns (3.8%; $F_{1,32} = 14.78$, $p < 0.01$). This interaction is

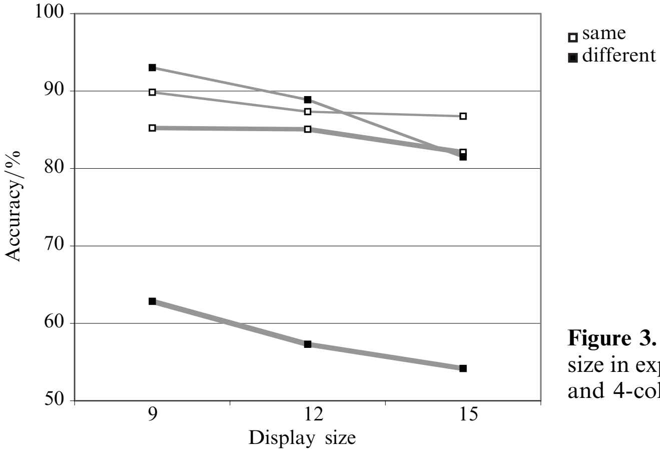

Figure 3. Accuracy as a function of display size in experiment 4, for 2-color (thin lines) and 4-color (thick lines) palettes.

predicted by the principle that color similarity aids detection of differences (Johnson et al 2009). There were also effects of display size ($F_{2,64} = 14.62$, $p < 0.001$, $\eta_p^2 = 0.10$) and an interaction of display size and response ($F_{2,64} = 5.04$, $p < 0.01$, $\eta_p^2 = 0.04$).

The advantage for 2-color palettes was generally consistent across the two color schemes (groups). The interaction of palette size and group was marginal ($F_{1,32} = 3.00$, $p = 0.09$) and overall performance was higher with yellow/green group (82.2%) than the red/purple group (76.8%) ($F_{1,32} = 8.13$, $p < 0.01$). The data for each color-scheme group are shown in table 1. There was no tendency for the advantage for 2-color palettes to increase with display size in this experiment ($F_{2,64} < 1$). Because there was only one 4-color palette for each group, further analyses of palette nested within palette-size were not conducted.

The capacity estimate was 11.7 stimulus units for 2-color palettes and 6.5 units for 4-color palettes ($t_{33} = 9.24$, $p < 0.001$, two tailed, SED $= 0.57$). The increase in stimulus capacity as palette size decreased from 4 to 2 was 44.9%.

Table 1. Percentage correct for each palette in experiment 4.

Palette size	Reds		Yellow/Greens	
2-color	87.1 (dark reds)	84.8 (light reds)	91.6 (yellows)	88 (greens)
4-color	67.6		74.6	
Difference	18.4		15.2	

6 General discussion

Humans rate similar-color palettes as more pleasant or harmonious than dissimilar-color palettes (eg Chevreul 1839/1987; Moon and Spencer 1944; Ou and Luo 2006; Schloss and Palmer 2011). Do such color relations influence the ease with which colored patterns are perceived and retained in mind? We found that they do, and that the critical color relation appears to be similarity of hue. High color-similarity palettes led to significantly higher performance, producing increases in estimated memory capacity, in terms of color units, of 26% to 45%. Thus, more similar (and more aesthetic) colors were easier to perceive and hold in VSTM. Our results suggest further that similarity of hue is more important than the harmony of the hues. Adding harmonious colors to a palette (experiment 4) led to a marked decrease in memory performance relative to a smaller (and more similar) color palette. However, we have not exhausted all possible approaches to color harmony. Indeed, we believe that further research on this topic can be interesting and useful.

The advantages for similar-color palettes have significant implications for graphic design in the real-world because of the ubiquitous nature of colors and colored designs. The results imply that people can integrate and retain much more information from displays of similar colors than from displays of dissimilar colors—as much as 45% more when palette size is reduced. The results extend the relation between aesthetics and mind, found previously for form, to the domains of color relations and memory performance.

However, the advantages of similar-color palettes may vary with the task. Note that the present VSTM task encourages the grouping together of colored units into larger (easier to retain) structures (eg Jiang et al 2000; Miller 1956; Sanocki et al 2010). Similarity of color may aid in the formation and maintenance of groups (see below). In other tasks, however, color dissimilarity might be helpful—eg in tasks that require filtering one color stimulus from others. In graphic design, color dissimilarity may be important for segregating selected information, creating color dominance, or creating tension.

The question why color similarity influences perceptual pleasure and VSTM is open and interesting. Explanations should address both dependent variables. Ultimately, explanations may be based in the neural coding of similar and dissimilar hues. However, the neural coding of color is an ongoing area of research, especially at higher cortical levels (eg Conway et al 2007; Stoughton and Conway 2008). Also, in our results the relation between individual ratings of harmony and VSTM performance were not high; this is an issue for further research.

Here, we consider two types of explanations of the similarity advantage in VSTM. Examination of the stimulus patterns (eg figures A1 and A5) suggests that there may be differences in the ease of grouping elements from similar and dissimilar palettes. There are at least two possible types of grouping differences. First, it may be easier with similar colors to integrate multicolor elements into larger groups; for example, yellows and greens may integrate together better than violets and greens, and identical colors are likely to group best of all. Grouping makes memory coding more efficient because there is a smaller number of groups than individual units, and more units can be held in memory as a result (Miller 1956). The gain in coding efficiency would be more for similar colors, and would be especially strong when there are fewer colors (experiment 4). A second type of grouping effect is that similar colors could reduce competition between separate groups; for example, yellow groups and green groups may coexist better than violet groups and green groups, perhaps because there is less inhibition between the color codes.(3) Ease of grouping and compatibility of groups could determine perceptual pleasure as well, with easily grouped units being viewed as more pleasant or harmonious.

A different type of explanation of the VSTM effect is based on the ease of detecting differences in similar and dissimilar contexts (Johnson et al 2009). Note that explanations in terms of ease of simultaneous color discrimination appear dubious in light of experiment 2. However, Lin and Luck (2009) recently reported an effect of stimulus similarity on change detection across delays that can be viewed as analogous to the present effects. Specifically, when a memory set contains similar colors and a single dissimilar color, detection of change is easier when the changed item is a variation of a similar color than when it is an (equally large) variation of the dissimilar color. The result occurs when the colors are presented simultaneously (but in separate locations) and when the memory colors are presented sequentially. The simultaneous presentation result is most analogous to the present experiments. The Johnson et al (2009) model

---

(3) Note that the likelihood was the same for similar and dissimilar palettes (experiments 1 and 3) for either the exact same colors to be adjacent elements within a pattern, or for members of row-pairs to be adjacent elements. Similarity effects resulted from grouping differences between row pairs—grouping must be stronger, or competition less, between similar row pairs than between dissimilar row pairs.

was based in part on the Lin and Luck (2009) results. As described earlier (section 2.3), the model's explanation depends on the balance of inhibitory effects between similar colors in VSTM on one hand, and between representation layers (for the first and second pattern) on the other hand. The similarity effect is produced when inhibition between similar colors results in reduced inhibitory input to the layer for the incoming test stimulus. The model predicts that color similarity effects will be greatest on different trials, and the present experiments were consistent with this prediction—marginal tendencies were obtained in experiments 1 and 3, and a large effect was obtained in experiment 4. One further challenge for this approach is explaining differences in perceptual pleasure for similar and dissimilar colors. Increased inhibition for similar colors does not seem a promising approach for explaining the great pleasure of their relations.

The relationship between perceptual pleasure, VSTM capacity, and color similarity could also be approached in terms of perceptual fluency. Previous research has demonstrated that observers are more positively disposed towards objects they have previously encountered (Zajonc 1971, 1997). While initial descriptions of the mere exposure effect hinged on a habituation of the orienting response, recent evidence is consistent with a perceptual-fluency hypothesis (Reber et al 1998). The perceptual-fluency hypothesis maintains that observers internally monitor their perceptual processes and experience affective responses associated with current processing demands. These emotional responses then inform evaluative judgments of the perceived object. When observers encounter an object that is easy to process, or perceptually fluent, they are more likely to regard that object positively. If something is easy to perceive, observers attribute their evaluation of the relative ease of perceiving the object to the object itself. In cases where the same object is presented repeatedly, subsequent presentations of the object are primed and easier to perceive. Similarly, when stimuli are presented under either clear or degraded conditions, observers rate the clearly present stimulus more positively. In the context of the present experiments, observers may rate similar-color palettes as more pleasant or harmonious specifically because they are easy to perceive and represent in VSTM (perhaps because of the similar color codes).

The present results provide further evidence that VSTM can hold fairly complex objects (Alvarez and Cavanagh 2004; Phillips 1974; Sanocki et al 2001, 2010). In the present experiments, VSTM capacities, in terms of number of color units, were as high as 12 stimulus units. Color relations influenced VSTM capacity, with similar colors increasing it by 26% to 45%. In the real-world, where humans must encode informational displays at high rates, this increase in encoding and memory maintenance may be very important.

Acknowledgments. Portions of these results were presented at the 9th annual meeting of the Vision Sciences Society, Naples, FL (May 2009). Thanks to Del Lindsey and Angela Brown for helpful comments, and to the members of the Vision Cognition Design Lab for their help running the experiments.

References

Alvarez G, Cavanagh P, 2004 "The capacity of visual short-term memory is set both by visual information load and by number of objects" *Psychological Science* **15** 106 – 111

Arnheim R, 1974 *Art and Visual Perception: A Psychology of the Creative Eye* (Berkeley, CA: University of California Press)

Baddeley A D, Hitch G J, 1974 "Working memory", in *The Psychology of Learning and Motivation* Ed. G H Bower (New York: Academic Press) pp 47 – 89

Biederman I, 1987 "Recognition-by-components: A theory of human image understanding" *Psychological Review* **94** 115 – 147

Chevreul M E, 1839/1987 *Principles of Harmony and Contrast of Colors and Their Applications to the Arts* (West Chester, PA: West Schiffer Publishing). Based on the first English edition of 1854 as translated from the first French edition, 1839

Chijiiwa H, 1987 *Color Harmony* (Rockport, MA: Rockport Publishers)

Conway B R, Moeller S, Tsao D Y, 2007 "Specialized color modules in macaque extrastriate cortex" *Neuron* **56** 560 – 573

Garner W R, 1974 *The Processing of Information and Structure* (Hillsdale, NJ: Lawrence Erlbaum Associates)

Granger G W, 1955 "Aesthetic measure applied to color harmony: an experimental test" *Journal of General Psychology* **52** 205 – 212

Hackett G, Schloss K B, Palmer S E, 2009 "The relationship between color and form in judgments of preference and harmony" *Journal of Vision* **9**(8):339, 339a [abstract], doi:10.1167/9.8.339

Jiang Y, Olson I R, Chun M M, 2000 "Organization of visual short-term memory" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **26** 683 – 702

Johnson J S, Spencer J P, Luck S J, Schoner G, 2009 "A dynamic neural field model of visual working memory and change detection" *Psychological Science* **20** 568 – 577

Lin P H, Luck S J, 2009 "The influence of similarity on visual working memory representations" *Visual Cognition* **17** 356 – 372

Logie R, 2003 "Spatial and visual working memory: A mental workspace" *Cognitive Vision* **42** 37 – 78

Luck S J, 2008 "Visual short-term memory", in *Visual Memory* Eds S J Luck, A Hollingworth (New York: Oxford University Press) pp 43 – 85

Miller G A, 1956 "The magical number seven plus or minus two: some limits on our capacity for processing information" *Psychological Review* **63** 81 – 97

Mollon J D, 2000 "Cherries among the leaves", in *Color Perception: Philosophical, Psychological, Artistic, and Computational Perspectives* Ed. S Davis (New York: Oxford University Press) pp 10 – 30

Moon P, Spencer D E, 1944 "Aesthetic measure applied to color harmony" *Journal of the Optical Society of America* **34** 234 – 242

Ou L, Luo M R, 2006 "A colour harmony model for two-colour combinations" *Color Research and Applications* **31** 191 – 204

Palmer S E, 1983 "The psychology of perceptual organization: A transformational approach", in *Human and Machine Vision* Eds J Beck, B Hope, A Rosenfeld (New York: Academic Press)

Pashler H, 1988 "Familiarity and visual change detection" *Perception & Psychophysics* **44** 369 – 378

Phillips W A, 1974 "On the distinction between sensory storage and short-term visual memory" *Perception & Psychophysics* **16** 283 – 290

Poggesi R M, Schloss K B, Palmer S E, 2009 "Preference for three-color combinations in varying proportions" *Journal of Vision* 9(8):332, 332a [abstract], doi:10.1167/9.8.332

Polzella D J, Montgomery D A, 1993 "Dimensions of color harmony" *Bulletin of the Psychonomic Society* **31** 423 – 425

Potter M, 1993 "Very short-term conceptual memory" *Memory & Cognition* **21** 156 – 161

Reber R, Winkielman P, Schwarz N, 1998 "Effects of perceptual fluency on affective judgments" *Psychological Science* **9** 45 – 48

Sanocki T, Sellers E, Mittelstadt J, 2001 "High capacity visual short term memory for layout" paper presented at the Annual Meeting of the Psychonomic Society Orlando, FL (November)

Sanocki T, Sellers E, Mittelstadt J, Sulman N, 2010 "How high is visual short term memory capacity for layout?" *Attention, Perception, & Psychophysics* **72** 1097 – 1109

Sanocki T, Sulman N, 2008 "Visual short term memory for location: Does objecthood matter?" *Journal of Vision* **8**(6):203, 203a [abstract], doi:10.1167/8.6.203

Schloss K B, Palmer S E, 2011 "Aesthetic response to color combinations: preference, harmony, and similarity" *Attention, Perception & Psychophysics* **73** 551 – 571

Smith V C, Pokorny J, 2003 "Color matching and color discrimination", in *The Science of Color* second edition, Ed. S K Shevell (Oxford: Elsevier) pp 103 – 148

Sloane P, 1989 *The Visual Nature of Color* (New York: Design Press)

Stoughton C M, Conway B R, 2008 "Neural basis for unique hues" *Current Biology* **18** R698 – R699

Vogel E K, Woodman G F, Luck S J, 2001 "Storage of features, conjunctions, and objects in visual working memory" *Journal of Experimental Psychology: Human Perception & Performance* **27** 92 – 114

Zajonc R B, 1971 "Attraction affiliation and attachment", in *Man and Beast: Comparative Social Behavior* Eds J E Eisenberg, W S Dillon (Washington, DC: Smithsonian Institution Press) pp 141 – 179

Zajonc R B, 1997 "Emotions", in *Handbook of Social Psychology* 4th edition, volume 1, Eds D Gilbert, S T Fiske, G Lindsey (New York: McGraw-Hill) pp 591 – 632

Zhang W, Luck S J, 2008 "Discrete fixed-resolution representations in visual working memory" *Nature* **453** 233 – 235

Appendix [In color online.]

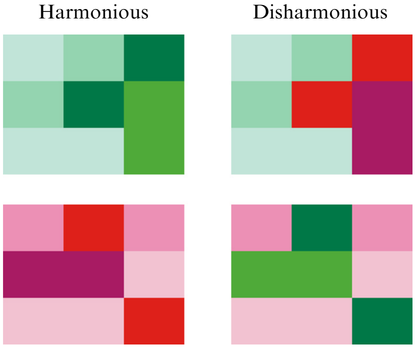

Figure A1. Examples of 8-element patterns, one from each of the four palettes in experiment 1. Note: patterns were randomly generated in the experiments; however, for comparison, these patterns are logically identical within rows.

Condition/ Color Category		CIE $L^*a^*b^*$ Coordinates		Harmony Rating
Harmonious				
green	🟩🟩	49.7, −47.1, 15.0	66.5, −48.7, 42.9	5.1
green (pale)	🟩🟩	89.8, −11.7, 0.87	82.7, −24.6, 7.7	
red	🟥🟥	52.7, 65.8, 53.1	44.6, 59.8, −5.8	6.7
red (pale)	🟥🟥	86.6, 16.0, −0.1	76.0, 33.5, −2.7	

Condition/ Color Category		CIE $L^*a^*b^*$ Coordinates		Harmony Rating
Disharmonious				
red	🟥🟥	52.7, 65.8, 53.1	44.6, 59.8, −5.8	1.9
green (pale)	🟩🟩	89.8, −11.7, 0.87	82.7, −24.6, 7.7	
green	🟩🟩	49.7, −47.1, 15.0	66.5, −48.7, 42.9	2.9
red (pale)	🟥🟥	86.6, 16.0, −0.1	76.0, 33.5, −2.7	

Figure A2. Palettes in experiment 1, with CIE $L^*a^*b^*$ coordinates in corresponding locations, and mean harmony rating (7-point scale).

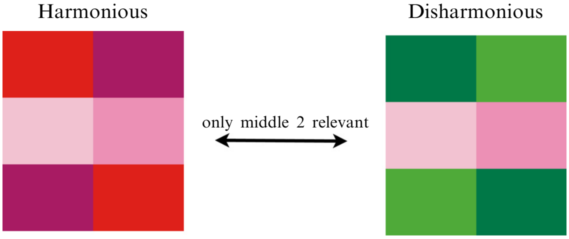

Figure A3. Two examples of stimulus patterns in experiment 2.

Condition/ Color Category		CIE $L^*a^*b^*$ Coordinates		Harmony Rating
Harmonious				
yellow	🟨🟨	95.1, −19.6, 71.9	96.5, −12.6, 22.4	72.8
green	🟩🟩	85.7, −78.7, 74.8	90.4, −47.7, 62.4	
blue	🟦🟪	42.9, 56.1, −88.2	71.6, 25.3, −40.7	79.8
violet	🟪🟪	57.0, 58.3, −53.5	74.7, 29.5, −28.4	

Condition/ Color Category		CIE $L^*a^*b^*$ Coordinates		Harmony Rating
Disharmonious				
violet	🟪🟪	57.0, 58.3, −53.5	74.7, 29.5, −28.4	44.9
green	🟩🟩	85.7, −78.7, 74.8	90.4, −47.7, 62.4	
blue	🟦🟪	42.9, 56.1, −88.2	71.6, 25.3, −40.7	33.8
yellow	🟨🟨	95.1, −19.6, 71.9	96.5, −12.6, 22.4	

Figure A4. Palettes in experiment 3, with CIE $L^*a^*b^*$ coordinates in corresponding locations, and mean harmony rating (100-point scale).

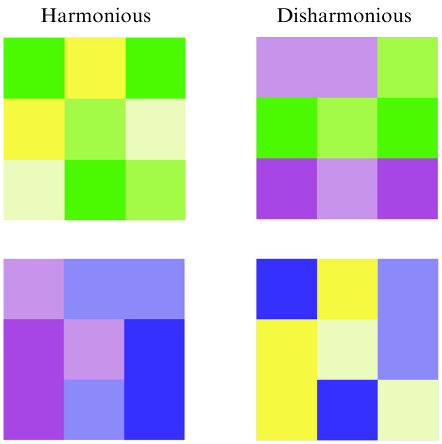

Figure A5. Examples of 9-element patterns, one from each of the four palettes in experiment 3.

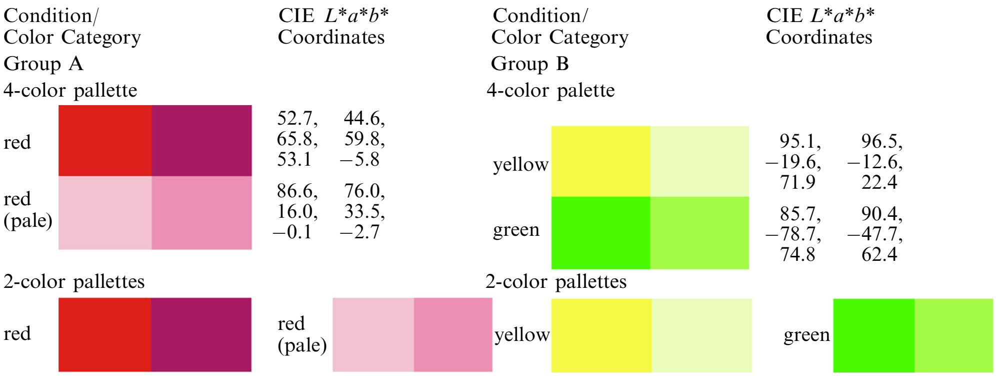

Figure A6. Color schemes and palettes in experiment 4, with CIE coordinates for the base 4-color palettes in corresponding locations.

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION
VOLUME 40 2011

www.perceptionweb.com