EFFECTS OF CATEGORIZATION ON COLOR PERCEPTION

Robert L. Goldstone
Indiana University

Abstract—Subjects were shown simple objects and were asked to reproduce the colors of the objects. Even though the objects remained on the screen while subjects reproduced the colors and the objects' shapes were irrelevant to the subjects' task, subjects' color perceptions were influenced by the shape category of an object. For example, objects that belonged to categories with redder objects were judged to be more red than identically colored objects belonging to another category. Further experiments showed that the object categories that subjects use, rather than being fixed, depend on the objects to which subjects are exposed.

The notion that experience and expectations can influence perception can be traced at least back to the Sapir-Whorf hypothesis (Whorf, 1941/1956) and the "New Look" movement (Bruner & Postman, 1949). This original work and its revivals (Niedenthal & Kitayama, 1994; von Hippel, Hawkins, & Narayan, 1994) stress that high-level cognitive processes do not simply operate on fixed perceptual inputs; high-level processes may also create lower level percepts. One notion from this literature is that our concepts and categories influence perception. The experiments reported here concerned the influence of learned categories on color contrast and assimilation effects. A contrast effect occurs when the dimension value ascribed to a stimulus is distorted away from its true value in the direction opposite to the other presented stimuli's dimension values. An assimilation effect occurs when the dimension value is distorted toward the values of other presented stimuli.

Contrast and assimilation effects may depend not just on the other set of stimuli in the experimental context, but also on similarities and categories that are formed among these stimuli. There is evidence that conceptual similarity influences contrast effects. In the Ebbinghaus illusion, the size of a central object appears to be smaller when it is surrounded by large, rather than small, objects. Coren and Enns (1993) have shown that this illusion is stronger when the central and surrounding objects belong to the same conceptual category (e.g., all the objects are dogs) than when the surrounding objects belong to a different category than the central object.

In the current experiments, subjects judged the color of objects that belonged to different shape or conceptual categories. Even though the object categories were irrelevant for the color judgment task, there was the possibility that these categories would exert an influence on color perception. Because of the experimental controls, if such an influence were found, it would have to be due to learned rather than preexperimental categories or category-to-color associations. Subjects assessed the color of an object by modifying a second identically shaped object until they thought it had the identical color. Contrast or assimilation effects were revealed by systematic misestimations of an object's color.

EXPERIMENT 1

To show an influence of shape categories on perception of an object’s hue, the ideal situation is to arrange two or more items from different categories to have objectively identical colors. The judgements of these equated objects can then be compared. In the two-categories condition of Experiment 1, six colored objects belonged to two categories: straight-edged letters and curved numerals. One of the letters had objectively the same hue as one of the numerals. The subjects’ task was simply to judge the hue of an object by adjusting an identically shaped object so that it possessed the same hue.

Method

Subjects
Eighty-five undergraduate students from Indiana University served as subjects in order to fulfill a course requirement. An additional 47 subjects were placed in the control condition.

Materials
Colored alphanumeric symbols were used as materials. The symbols T, E, L, 8, 9, and 6 were rendered on Macintosh II/1 screens in the Geneva type font. Each symbol was approximately 5.2 cm high by 4.6 cm wide. The luminance of all of the shapes was 27.6 cd/m². As represented in Figure 1, each symbol was given one of five hues along a continuum from red to violet. Going from the most red to the most violet hue, the 1976 CIE (Commission Internationale de L’Eclairage) values for the five hues were as follows: \( u' = 0.3291, v' = 0.4603; u' = 0.3017, v' = 0.4097; u' = 0.2788, v' = 0.3676; u' = 0.2631, v' = 0.3376; \) and \( u' = 0.2480, v' = 0.3134 \). These values were chosen so that subjective differences between adjacent hues were approximately equal (Goldstone, 1994), and so that all of the hues were noticeably different from each other. If the reference point is a black body source that produces equal energy at all wavelengths \( (u' = 0.2009, v' = 0.4609) \), the five hues can be assigned the following wavelengths (in nanometers): 511, 527, 541, 554, and 565.

Procedure
On each trial, a colored symbol was displayed in the upper left corner of the computer screen. A second black token of the same symbol was displaced 14 cm from the first symbol, toward the lower right corner of the screen. A horizontal 27-cm black line was presented at the bottom of the screen. The left end of the line was labeled "Red" and the right end was labeled "Violet." By moving a cursor along this line with a mouse, subjects

Address correspondence to Robert Goldstone, Psychology Department, Indiana University, Bloomington, IN 47405; e-mail: rgoldsto@ucs.indiana.edu.
were able to alter the appearance of the second symbol immediately. Subjects were instructed to adjust the second symbol’s color so as to match the first symbol’s color. When the two tokens were judged to have identical colors, subjects registered their judgment by pressing a button on top of the mouse. In order to ensure that judgments were based on the actual color of the symbol rather than absolute position on the hue scale, the red-to-violet scale was offset by a random physical distance between 2 and 8 cm from the left edge of the horizontal scale line.

Figure 1 shows only one of the stimuli counterbalancings used. For half of the subjects, E, L, and T were given red hues; for the other half of subjects, 6, 8, and 9 were given red hues. For each subject, the assignment of particular hues to the three numerals and the three letters was randomized. In the discussion of results, each of the six symbols in Figure 1 refers to any items that had the same role as that symbol. For example, “T” is used to refer to the reddest symbol, whatever that symbol was for a subject. In all cases, one of the letters and one of the numerals shared the central hue value in Figure 1.

Subjects saw each of the six symbols repeated 36 times. After subjects finished estimating the hue of a symbol, the screen was erased for 1 s, and then the next trial commenced. The entire experiment took approximately 35 min to complete.

An additional control condition was run to verify that misestimations were caused by letter and digit categories rather than simply the relative hues of the six symbols. In this control condition, the same symbol, randomly selected from the set of six, was displayed on each trial, but on each trial was randomly given one of the five hues used in the two-categories condition. The middle hue was used twice as frequently as the other four hues, as was the case in the two-categories condition.

Results and Discussion

The results of central interest concern how subjects’ estimates for hues departed from the symbols’ actual hues. Not all of the systematic misestimations were necessarily due to the category membership (letters vs. numerals) of symbols. Several perceptual factors could create small systematic distortions in perceived hues (Abramov & Gordon, 1994), such as a bias to see a red-orange hue as more prototypically red than it actually is. However, if we restrict our attention to L and 8 in Figure 1, and if there is a different pattern of judgments between them, then the difference cannot be attributable to perceptual properties of particular hues: L and 8 have the same hue.

As shown in Figure 2, the mean overestimations, in nanometers, for T, E, L, 8, 9, and 6 were –2.24, 0.16, 0.56, 1.84, 3.20, and 2.00, respectively. Positive overestimations indicate that the hue was judged to be more violet than it actually was; underestimations (negative overestimations) indicate the hue was judged to be redder than it actually was. A planned t test indicated that 8 was overestimated to a greater extent than was L, t(84) = 10.3, p < .01. The amount of misestimation for each of the symbols, with the exception of E, significantly departed from 0.0, t(84) > 2.8, p < .05. Overestimation differences of 0.42 are significant by Fisher’s post hoc probabilistic least significant difference (PLSD), p < .05.

In general, the estimates for the symbols that belonged to the more violet category (8, 9, and 6 in Fig. 2) were displaced to the violet end of the scale, relative to symbols that belonged to the redder category (T, E, and L). Although 9, 6, E, and T do not have controls in the two-category condition, estimates for these symbols from the two-category condition can be compared with the estimates obtained from the one-category control condition. Relative to their controls, T was underestimated (t(130) = 8.7, p < .01), E was underestimated (t(130) = 2.4, p < .05), 9 was overestimated (t(130) = 2.7, p < .05), and 6 was insignificantly underestimated (t(130) = 1.4, p > .1). The pattern of results for T, E, and 9 is not compatible with a model that distorts an item’s perception in the direction of the average dimension value for the category. Such a model would predict that T would be overestimated (toward the average hue for letters or symbols in general) and E and 9 would not be systematically misestimated. Thus, the influence of categorization seems to implicate a polarization or caricaturization (Goldstone, 1993).
Color Perception

Fig. 2. Mean estimates of hue from Experiment 1. Positive overestimations indicate that a hue was judged to be more violet than it actually was. Negative overestimations (underestimations) indicate a hue was judged to be redder than it actually was. Error bars show the standard deviations associated with each overestimation.

process that can be stated as follows: “If X belongs to a category that has a large (or small) value on a dimension relative to other categories, then distort X’s perceived dimension value to make it appear larger (or smaller).”

The results from the control condition suggest that T’s underestimation and 6’s overestimation were partially due to anchor contrast effects that moved endpoint items in the direction opposite to the other presented items (M.C. King, 1981; Lane, 1965). However, the differences between the two-categories condition and the one-category control indicate that category-insensitive endpoint effects are not sufficient to explain the results. Overall, Experiment 1 shows a category-level influence on symbols that is compatible with a polarization process that makes objects that belong to a category that contains relatively red (violet) objects seem even more red (violet).

EXPERIMENT 2

Experiment 2 further tested the influence of category-level information on hue perception by exploring the nature of the categories that influence perception. Because Experiment 1 included no explicit categorization task and therefore no feedback about what categories were correct, the results might be interpreted as reflecting fixed constraints on what categories will influence perception. For example, if subjects sort objects into numerals and letters, one could argue that these categories must be automatic and context-independent ways of sorting the symbols.

Although no category feedback was provided, category-level information was provided by the particular items shown to subjects (Clapper & Bower, 1994; Rumelhart & Zipser, 1985). If a set contains 8, 6, E, and T, then subjects might spontaneously create categories for numerals and letters, but if squares and triangles are also shown, then subjects might create categories of symbols and shapes. Consequently, Experiment 2 explored the possibility that fixed categories do not determine perceptual distortion, but rather subjects spontaneously create context-dependent categories that influence hue perception.

As shown in Figure 3, within the low-similarity set of stimuli for Experiment 2, there appear to be two categories of shapes: five-sided polygons and two-lined branches. If subjects are sensitive to these categories, then assimilation within these categories would be expected, as was found in Experiment 1. That is, Object D may be assimilated toward the other polygon in the low-similarity set. However, in the high-similarity set, the identical polygon may not be judged as belonging to the same category. In this set, all of the shapes are five-sided polygons, and

Fig. 3. Four conditions used in Experiment 2. The shading of a shape indicates its proportion of red, relative to violet, hue. Shapes A, B, C, and D are identical. The high-similarity set is identical to Control 1, with the addition of one upward-pointing polygon. The low-similarity set is identical to Control 2, with the addition of the same upward-pointing polygon.
consequently this description lacks diagnosticity. Instead, the naturally constructed categories may be upward-pointing prongs and right-pointing prongs. If these categories are created, then C would not be assimilated toward the other polygons. Consequently, although C and D are identically shaped and colored, D may be judged to be redder than C. In short, depending on the entire set of shapes in a context, the same two shapes may or may not be placed in the same category. The pattern of hue distortions can reveal the nature of the implicitly formed categories.

**Method**

**Subjects**

Sixty-three undergraduate students from Indiana University were evenly assigned in a pseudorandom manner to the three presentation conditions.

**Materials**

Colored shapes, rendered on Macintosh I1SI computers and shown in Figure 3, were used as materials. The dominant wavelengths for the three hues that were used were 516, 541, and 565 nm. The CIE coordinates for the 516-nm shapes were \( u' = .308 \) and \( v' = .222 \), for the 541-nm shapes were \( u' = .296 \) and \( v' = .399 \), and for the 565-nm shapes were \( u' = .259 \) and \( v' = .300 \).

The shapes were designed to fall into two shape categories: five-sided polygons and two conjoined thick lines. In addition, Shape A (B, C, and D are identical to A) was designed to be clearly distinguishable from the other polygons. A’s prongs face to the right whereas the other polygons’ prongs face upward.

**Procedure**

The experiment’s procedure was essentially the same as used in Experiment 1. While a colored shape was displayed, subjects modified a replica of the shape so that it possessed the same hue. The same scale and trial randomizations were used.

Three separate groups of subjects saw three different sets of shapes (combining the two control conditions into one set). The three shape sets are shown in Figure 3. In the high-similarity set, all the shapes are polygons; in the low-similarity set, two of the shapes are polygons and two are composed of two thick lines. Subjects saw 36 repetitions of each of the three or four shapes in a condition. On each presentation, the shape possessed the hue indicated by its position on the hue dimension in Figure 3. The experiment took approximately 30 min for subjects to complete.

**Results and Discussion**

The two control conditions produced highly similar patterns of misesestimation, and, consequently, the results from Objects A and B were combined. The cleanest comparisons are between shapes that have identical hues. The experiment was explicitly designed to manipulate perceptions of the centrally hue shape (A, B, C, and D in Fig. 3). The distribution of hue overestimations for these shapes is given in Figure 4. For these shapes, the set context had a significant influence on hue overestimation,

Fig. 4. Mean estimates of hue from Experiment 2. The set context influenced the hue perception of objects that had identical hues and shapes. Although Objects A, B, C, and D were identically shaped and hued, as shown in Figure 3, they were presented in different contexts, and consequently subjects’ estimates of their hues were different. Positive overestimations indicate that a hue was judged to be more violet than it actually was. Negative overestimations (underestimations) indicate a hue was judged to be redder than it actually was. The vertical axis represents the percentage of responses that involved a particular degree of overestimation.

\[ F(2, 40) = 16.3, p < .01, \] and each of the average overestimations was significantly different from the others, \( F(20) > 2.0, p < .05 \). As Figure 4 shows, relative to B (second control condition), D was underestimated (low-similarity set). However, C (high-similarity set) was overestimated relative to A. These results are consistent with the two polygons belonging to the same category in the low-similarity set, but the same two shapes belonging to different categories in the high-similarity set, in which the added upward-pointing polygon is in the context of two other upward-pointing polygons. The pattern of results in Figure 4 indicates both a repelling force of the added upward-pointing polygon on C (relative to A’s perceived hue), when the polygons belonged to different categories, and an attracting force of the added upward-pointing polygon on D (relative to B), when the polygons were likely to be categorized together.

**GENERAL DISCUSSION**

Experiments 1 and 2 demonstrate that subjects’ hue perceptions are distorted by the judged objects’ shape categories. The results indicate an automatic process of shape categorization that exerts a perceptual influence on hue. The categorization is automatic in the sense that shape categorization is not required in order to accomplish the task, and yet still is naturally performed in the course of a color perception task. Task counterbalancing and same-hue controls indicate that explanations for perceptual distortions must taken into account the shape category of judged objects. Experiment 1 shows that shape categories can influence hue perception despite their lack of preexperimental association to hue, and Experiment 2 shows that the shape categories that guide hue perception are defined by the experimental context rather than being preset.
Color Perception

Fixed and Created Categories

Previous research has shown that the color appearance of an object is influenced by the context (Bruner, Postman, & Rodriguez, 1951; Delk & Fillenbaum, 1965; Harper, 1953; White & Montgomery, 1976). Among other results, these experiments show that the color of an object is distorted toward the color suggested by the object’s shape. A color patch placed on a banana-shaped figure is matched to a truer yellow than the same color on a strawberry shape.

Other research, including the current experiments, shows that new contexts can influence perception of a dimension even though they have no preexperimental associations to the dimensional (Manis, Nelson, & Shedler, 1988; Marks, 1992; Marks & Warner, 1991; Wedell, in press). Although many of these experiments have shown contrast effects rather than the assimilation effects reported here, they do argue for the creation of subcontexts within the larger experimental context. For example, Marks showed that under conditions in which a 500-Hz tone was generally loud and a 2,500-Hz tone was generally soft, a 2,500-Hz tone was judged to be much louder than an equally loud 500-Hz tone.

The current experiments further indicate that contexts cannot simply be based on single dimension values or preset categories. The results of Experiment 2 suggest that two objects were placed in the same category when other objects were highly dissimilar to the two objects, but the same two objects were placed in different categories when other objects were similar. Some researchers have found evidence for contexts that are defined by critical boundaries on physical dimensions (Marks & Warner, 1991): Events that differ by more than x units are placed in different contexts. The current experiments indicate that, in some circumstances, the critical boundary x must be flexibly tuned to the stimulus set rather than being a single fixed value.

The context-defining categories seem not to be preset, but to be learned during the experiment. Category learning can affect the perceptual discrimination of color attributes, such that attributes that are diagnostic for categorization become perceptually sensitized (Goldstone, 1994). However, the current experiments show that category learning can also proceed without explicit categorization feedback (cf. Clapper & Bower, 1994). The technique offers a potential tool that can reveal people’s implicit categories and that is relatively free of task demands and the problem-solving stance typical of many category-learning paradigms.

Contrast or Assimilation

In Experiment 1, objects with the same hues were perceptually distorted in the direction of the object in their context. This effect is in the opposite direction of what Wedell (in press) and Marks (1992) observed. Other researchers, however, have found results in the same direction as the current results (Manis, Biernat, & Nelson, 1991; Tafjel & Wilkes, 1963). The discrepancy may be due to factors such as object extremity, object ambiguity, sequence effects, or resource demands that have been shown to mediate whether contrast or assimilation effects are found (Foti & Hauenstein, 1993; Herr, Sherman, & Fazio, 1983; Lockhead & King, 1983; Manis et al., 1988; Ward, 1990; Wedell, 1990). In particular, one possibility is suggested by Biernat, Manis, and Nelson’s (1991) result that objective scales are more likely to produce assimilation effects than are subjective rating scales. In Experiments 1 and 2, precautions were taken to ensure that subjects were making objective hue judgments rather than using an arbitrary scale.

Experiment 1 may reflect either a contrast or an assimilation effect. For example, L in Figure 1 is underestimated relative to 8. This result may be due to L being assimilated to the other letters, or to L being contrasted away from the numerals. Experiment 2 teased apart these two possible contributions to perceptual distortions, and suggests an influence of both processes—assimilation within the immediate category and contrast from dissimilar categories. Assimilation within the immediate category is suggested by comparing the low-similarity set with the controls in Figure 3. When an object likely to be placed in the same category as D is introduced, then estimates of D move toward that object. Contrast from competing categories is suggested by comparing the high-similarity set to the controls. The introduction of an object that hypothetically belongs to a different category moves C’s estimates away from the object.

One line of research (D.L. King, 1988; Lockhead, 1988) argues that assimilation effects are found when two events are perceived as a single whole, and contrast effects are found when two separable events are perceived. The current results are consistent with this claim. Objects that “go together” assimilate to each other, and objects that are psychologically separated are contrasted. Objects may go together because they are close in time (Lockhead & King, 1983; Ward, 1990), close in space (Coren & Gurgus, 1980), or conceptually similar (Coren & Enns, 1993). The current manipulation of experimentally presented categories may simply be another way of altering the grouping of objects. Subjects seem to create categories that will make a coarse split between all of the presented objects. Within a coarse category, objects seem to be assimilated. Objects seem to be contrasted away from objects that fall on the opposite side of the coarse category split.

The Locus of Categorization’s Influence

The nature of the color judgment task eliminated several potential explanations that do not posit a perceptual effect of object categories. Influences of memory encoding should have been minimal. Although there is evidence that expectations can bias people’s memories for colors (Belli, 1988), in the current experiments, only very short-term memory was required. The displayed object and the manipulated object were viewed at the same time, and it took very little time to scan from one object to the other. Furthermore, because the subjects’ task was to match two objects’ colors, there were few biases associated with numeric rating scales.

More generally, the results may be compatible with either perceptual encoding or response selection accounts of categorization’s influence on judgments. Other research has suggested that contrast and assimilation effects occur early in the encoding stage of processing (Sullivan & Wyer, 1980; Wedell, 1994).
Although the current results concur with these findings to the extent that the hue judgments required of subjects were direct and simple, the results also indicate that categorization processes operate before contrast and assimilation processes. Categorization, or the creation of separate shape-based contexts within the overall experimental context, is a necessary precursor for the obtained perceptual distortions to occur.

Possible Models

One type of model for the type of perceptual distortions of $8$ and $L$ found in Experiment 1 maintains that judgments are combinations of particular item information and more general category-level information (Hellstrom, 1985). For example, Huttonlocher, Hedges, and Duncan (1991) modeled judgments about the spatial location of an object by combining evidence from the specific object’s location and from the general spatial category of the object’s location. In Experiment 1, there was a trend (significant for one of the two objects) for the endpoint objects to be contrasted away from their category means. Although incompatible with a model that simply averages specific and category-level means, this result is consistent with a model that treats category-level information as relational. For example, in Experiment 1, the letter category may have been represented, in part, as “red, relative to the numerals” (Goldstone, 1993). This relational, rather than absolute, category-level information could distort even the reddest letter so that it would appear redder than it was. The result is also consistent with Krueger and Rothbart’s (1990) evidence that people distort item representations in order to make the items’ categories highly differentiated and separated from each other.

In order to make hue distortions contingent on categories that are learned during the experiment rather than preset, it may be useful to borrow techniques from unsupervised connectionist networks (Rumelhart & Zipser, 1985). Such networks can create categories on the basis of the statistical properties of the presented objects even when no explicit categorization feedback is provided. These models are able to create coarse categories when stimuli vary widely, and finer categories when all the stimuli are highly similar. A model of the hue distortions obtained in the current experiments should be sensitive to object categories, and should allow these categories to depend on the similarity relations between the presented objects.

Acknowledgments—I would like to thank Gregory Lockhead, Doug Medin, Paula Niedenthal, Robert Nosofsky, Richard Shiffrin, Jim Sherman, and Doug Wedell for many useful comments. The research was supported by National Science Foundation Grant SBR-9409232.

REFERENCES


Color Perception


(RECEIVED 7/18/94; REVISION ACCEPTED 4/28/95)