Cross-modal correspondence between vision and olfaction: The color of smells

AVERY N. GILBERT, ROBYN MARTIN, and SARAH E. KEMP
Givaudan-Roure Fragrances, New Jersey

Cross-modal sensory correspondences between vision and audition have been well described, but those between vision and olfaction have not. In Experiment 1, a method previously used to relate color names, mood names, and line elements was replicated and extended to describe odors by color. Significant color characterizations were found for all 20 test odors. Test-retest correlations showed color-odor correspondences to be as stable as nonodor measures after 2 years. In Experiment 2, new subjects matched Munsell color chips to the test odors. Thirteen odors had characteristic hues; there was significant variation in chroma and value. The selelcted Munsell hues corresponded to the color names endorsed in Experiment 1. Together, these experiments suggest the existence of robust correspondences between vision and olfaction.

Although the relationship between audition and vision has been extensively documented (Marks, 1975), a cross-modal linkage between the olfactory and visual senses is seldom mentioned in the literature. In one of the few studies available, Martin (1909) found that olfactory impressions were among the sensory experiences reported by subjects when shown reproductions of paintings. Odor experience was evoked in some subjects by stimulation of other senses following mescaline ingestion (Simpson & McKellar, 1955). Odors have been anecdotally reported to evoke synesthetic experiences in certain subjects (Cytowic, 1993). The paucity of studies appears to throw doubt on the idea that olfactory-visual correspondences exist. This impression is reinforced by Marks (1978) in his summary of the field:

> Synesthesias that involve the so-called minor or lower senses of taste, smell, touch, temperature, and pain stand in marked contrast to synesthesias of sight and sound. Whereas colored hearing synesthesia is dominated by regular and systematic correlations between visual and auditory dimensions of sensory experience, other forms of synesthesia are generally much more erratic and idiosyncratic, in that they rarely reveal common patterns or dimensions. (pp. 98–99)

AMERICAN JOURNAL OF PSYCHOLOGY
Fall 1996, Vol. 109, No. 3, pp. 335–351

There are other reasons to regard color-odor linkages as unlikely. A widely held view is that odor associations, preferences, and memories are very strongly encoded and that they are derived from idiosyncratic personal experience (Ehrlichman & Bastone, 1992). If this "Proustian" view is correct, we would expect little consensus among subjects regarding the color connotations of odor stimuli.

Another well-documented feature of odor perception is the difficulty people have in naming odors, even familiar ones (Cain, 1979). This so-called "tip-of-the-nose phenomenon" (Lawless & Engen, 1977) appears to be due to difficulty in accessing semantic labels for odors. It has been proposed that cross-modal correspondences occur through the mediation of semantic labeling (see Melara, 1989). If so, the existence of the tip-of-the-nose phenomenon implies that odor labels will not be readily available for use in generating color characterizations, further reducing the likelihood of finding a consistent color-odor linkage.

Despite these negative indications, it has been shown that color cues can bias odor judgments (e.g., Engen, 1972; Zellner & Kautz, 1990). The appropriateness of the color-odor pairing determines whether the bias aids or impairs performance (Burghardt, 1977; Davis, 1981; Zellner, Bartoli, & Eckard, 1991). Appropriate colors increase and inappropriate colors reduce the accuracy of odor identification. This implies that there are "correct" colors for smells. To date, the appropriate pairings have been assigned by experimenters, sometimes on the basis of opinion surveys (Zellner et al., 1991), but not on direct sensory evaluation.

These conflicting findings indicate the need for a quantitative basis on which to judge whether color-odor correspondences exist, whether they are idiosyncratic or consistent across subjects, and whether they are consistent over time. The two experiments reported here establish an empirical foundation on which to address these issues. The results, from both a color name task and a color matching task, suggest that odors can evoke characteristic hues.

EXPERIMENT 1

Rader (1979) and Rader and Tellegen (1987) developed a questionnaire-based protocol to quantify the synesthetic relations between color, mood, and line elements. Experiment 1 was designed to replicate Rader and Tellegen's results for color, mood, and lines and to extend the technique to odor and color relations. In addition, a subset of subjects was retested after 2 years to examine the consistency of results over time.

METHOD

Subjects

Subjects were 38 men and 56 women (18–40 years, $M = 29 \pm 7$ years). They were recruited through newspaper advertisements and screened by self-report for normal sense of smell, normal nasal breathing, and absence of active head cold, sinus infection, or allergy. They were also screened for color vision anomalies using the Ishihara test (1990). Subjects whose color vision was anomalous according to the Ishihara criterion (i.e., those reading fewer than 10 of Plates 1 to 11) were dropped from further analysis. Subjects provided their informed consent and were paid for participating.

Synesthesia questionnaires

Subjects were given the three synesthesia questionnaires described by Rader and Tellegen (1987). In the basic questionnaire format, the subject was given a stimulus (a mood name or a line drawing), a printed list of descriptors, and five "points." The subject distributed the points across the descriptors so as to best describe the stimulus. Fractional points were not permitted. Descriptors were listed in the same order for every stimulus. Stimuli were randomized within each questionnaire. Questionnaires were presented in the order: mood-color, line-mood, and line-color.

In the Mood-Color Questionnaire, the stimuli were 11 mood names (angry, tired, proud, peaceful, surprised, sad, excited, tender, ashamed, joyful, and scared) which the subject described by allocating five points across 11 color names (red, orange, yellow, green, blue, purple, brown, white, pink, gray, and black). The selection of these mood and color descriptors by Rader and Tellegen (1987) was based, in turn, on protocols used in experiments by previous investigators. The descriptors were adopted here to maintain consistency with the existing literature.

The Line-Mood Questionnaire used the Poffenberger and Barrows (1924) lines as stimuli. These 18 lines consist of two styles (angular and undulating), three spatial frequencies (slow, medium, and fast), and three inclinations (horizontal, tilted up, and tilted down). The subject described the emotional quality of the line stimuli using the five-point allocation procedure and the previous 11 mood names.

In the Line-Color Questionnaire, the stimuli were the Poffenberger and Barrows (1924) lines, which the subject described using the previous 11 color names.

Odor-Color Test

To characterize the cross-modal relation, if any, between odor and color, the Rader and Tellegen (1987) point allocation procedure was used in a novel odor-color task. The subject smelled an odor sample and described it by distributing five points across the 11 color names used previously.

Odor stimuli

Twenty test odorants (Table 1) were chosen from the fragrance materials used in commercial perfumery. Stimuli were selected to be representative of different fragrance types. They included essential oils extracted from botanical sources as well as synthetic aroma compounds. Some of the test odors may have been familiar to the general public, but others were unlikely to have been experienced outside of the fragrance industry.

Fragrance materials were prepared as solutions in diethyl phthalate, which dissolves a wide range of odorants and has minimal odor itself. Magnitude estimation of perceived odor intensity was used in pilot testing to adjust the stimuli to an intensity standard of 20% olibanum oil. Ten grams of odor solution were presented in 1 oz. (approximately 29 ml) opaque white high density polyethylene bottles labeled with random three-digit codes. The stimulus set was replaced with fresh solutions dàily.

Stimulus odors were presented in a single test session with order of presentation randomized for each subject. A minimum 20-s interval was maintained between odor presentations. Tests were conducted in a neutral-colored room with fluorescent lighting.

Table 1. Odor stimuli and concentrations (w/w in diethyl phthalate) used in Experiments 1 and 2

Substance	Concentration (%)
Aldehyde C-16	60
Bergamot oil	100
Birch tar oil	0.1
Caramel lactone	0.005
Cinnamic aldehyde	20
Civet artificial	5
2-Ethyl fenchol	10
Galbanum oil	1
Jasmine absolute	50
Lavender oil	100
Lilial	100
Methyl anthranilate	40
Methyl salicylate	10
Neroli oil	100
Olibanum oil	20
Patchouly oil	100
Pine oil	10
Rosalva	15
Star anise oil	20
Tarragon oil	10

Analysis

Neither Rader (1979) nor Rader and Tellegen (1987) offered a probability-based criterion by which to evaluate the mean scores resulting from the point assignment technique. To address this, we used the Friedman two-way analysis of variance (ANOVA) by ranks. Color scores for an odor were rank-ordered for each subject. Comparison of rankings across subjects permits a test of the null hypothesis that subjects assigned points to descriptors randomly. Because many tied observations (zeroes) are produced by the Rader and Tellegen point assignment procedure, the Friedman tests were repeated using O'Mahony's (1986) correction for ties. The tie correction procedure yields less conservative estimates of significance.

Call-back study.

Fifteen subjects (7 men, 8 women) took part in a call-back study. These subjects were those available from 10 men and 10 women contacted at random from a list of original participants. The minimum test-retest interval was 2 years and 2 months (range 767–798 days, $M = 784 \pm 10$ days). The call-back study was in all essential respects a repeat of Experiment 1. Subjects completed Questionnaires 1–4 in sequence, but with rerandomized stimulus order. The mean results of the call-back sample were compared with the means of the same group of subjects in the initial test. Mean descriptor scores from the initial and repeat tests were paired and used to calculate a correlation coefficient (r) for each stimulus item. These r values were transformed to Fisher's Z values, and used to calculate a mean correlation coefficient for each questionnaire.

RESULTS

Synesthesia questionnaires

Color descriptions of moods followed those outlined by Rader and Tellegen (1987). For example, joyful was characterized as yellow, surprised as red and yellow, peaceful as blue, tired as gray, scared as black, and sad as gray and black. In general, positive moods were characterized as red, blue, yellow, orange, and pink, and negative moods as brown, gray, and black. All 11 moods produced significant differences across colors on the Friedman two-way ANOVA ($p < .01$). Corrected for tied observations, all results were significant ($p < .01$).

Mood descriptions of lines were similar to those described by Rader (1979) and Rader and Tellegen (1987) and resembled the earlier results, obtained by a different protocol, of Poffenberger and Barrows (1924). For example, fast angular lines were characterized as predominantly angry, and were rated as progressively more scared and less excited as the left-to-right orientation changed from up to level to down. Seventeen of 18 line stimuli produced significant differences across moods on the Friedman two-way ANOVA ($p < .01$). Corrected for tied observations, all 18 results were significant ($p < .01$).

Color descriptions of lines were similar to those obtained by Rader (1979). For example, undulating lines in the level orientation were characterized as blue, and as less so with increasing spatial frequency. For angular lines, the red component decreased with decreasing spatial frequency, regardless of line orientation. Twelve of the 18 line stimuli produced significant differences across colors on the Friedman two-way ANOVA ($p < .01$). When corrected for tied observations, 16 of the 18 line stimuli produced significant results ($p < .01$).

Color-Odor Test

All but three odors (lavender, jasmine, and star anise) produced significant results on the Friedman two-way ANOVA ($p < .01$). Corrected for ties, all 20 odors showed significant results ($p < .01$). Thus, the color score profiles for the majority of odors were significantly different from those expected under the null hypothesis, namely a random assignment of points to colors (Figure 1). Heterogeneity in color score profiles was evident among the different odors. For example, a single, heavily endorsed color characterized some odors (e.g., brown for caramel lactone), whereas a range of colors was heavily endorsed for others (e.g., brown, white, gray, and black for civet). The highest scoring color names varied from odor to odor, and included hues as different as red, yellow, green, brown, and black.

High-scoring color names were widely endorsed. They do not appear to have been biased by high point assignment from relatively few subjects. Color scores in the top quartile of the distribution (*M*s > .60) were endorsed by an average of 36 ± 8% of subjects; means in the second, third, and fourth quartiles were endorsed by 22 ± 4%, 16 ± 3%, and 8 ± 3% of subjects, respectively.

Call-back study

For the experimental tasks of Rader and Tellegen (1987), mean test-retest correlations were highest for moods described with color terms ($r = .84$), lowest for lines described with color terms ($r = .57$), and intermediate for lines described by mood terms ($r = .78$). All three tasks had individual items with high correlations ($r \geq .93$). Variation in the number of near-zero item-wise correlations accounted for test-retest differences among the three tasks.

The mean test-retest correlation for odors described with color terms was $r = .53$. This was similar to the r for lines described by color terms. It was lower than that for moods described by color terms, or lines described by mood terms. Nine individual odors had significant test-retest

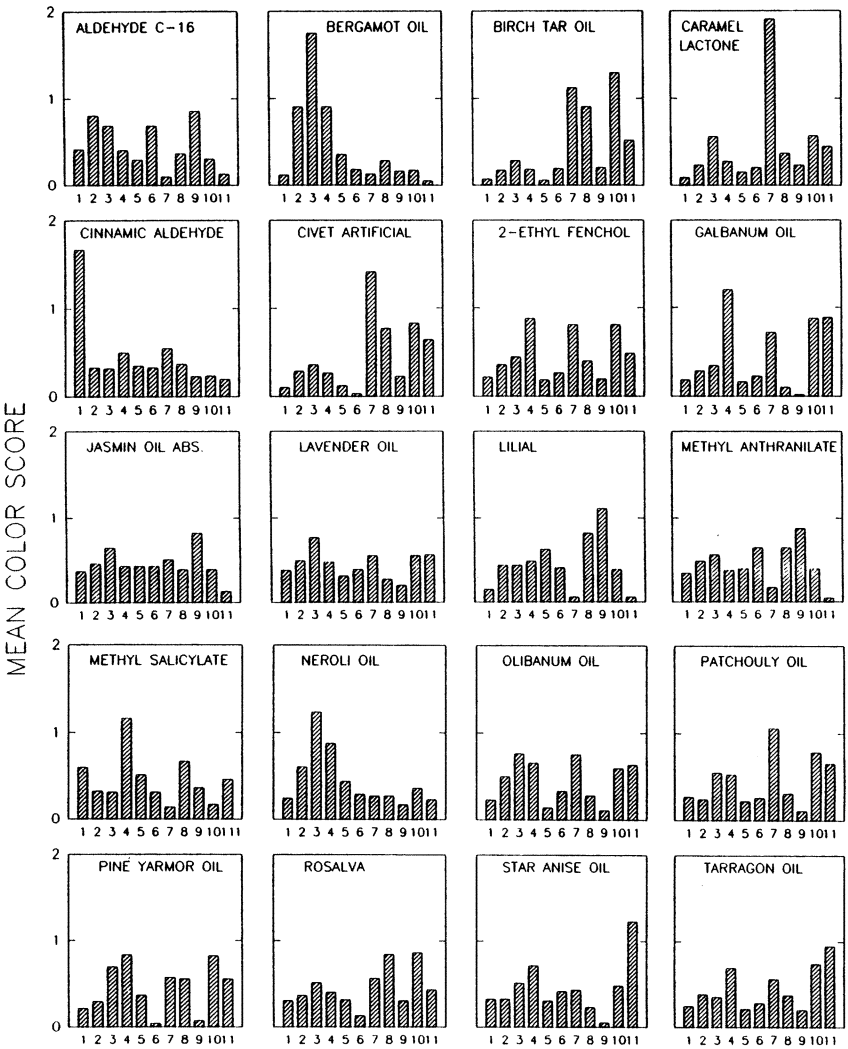

Figure 1. Mean color scores for the odor stimuli in Experiment 1 (1 = red, 2 = orange, 3 = yellow, 4 = green, 5 = blue, 6 = purple, 7 = brown, 8 = white, 9 = pink, 10 = gray, 11 = black)

correlations ($p < .05$ at $r \geq .60$, $df = 9$). For seven odors, the highest scoring color term was the same on the initial and follow-up tests; for one odor, the highest scoring color term exchanged ranks with the next highest.

DISCUSSION

Rader and Tellegen's (1987) point-allocation procedure proved to be replicable and robust. We obtained results similar to theirs on the tasks relating color names, mood names, and line elements. Our extension of this procedure to olfactory stimuli was also successful. The evidence suggests that odors are nonrandomly characterized by colors, or at least by color names. Compared with the three nonolfactory cross-modal tasks, the stimuli in the color-odor test produced results of similar magnitude and statistical significance.

Nonrandom color-odor correspondences were found despite expectations to the contrary derived from the literature on olfactory perception. The Proustian view holds that odor connotations are the result of powerful and idiosyncratic experiential factors. If the Proustian effect was operative in Experiment 1, there would have been large intersubject variability in the data, and few significant departures from the null hypothesis. The fact that we found a significant color consensus for many test odors leads us to suspect that the Proustian view has been overemphasized in the literature (Ehrlichman & Bastone, 1992; Gilbert & Kare, 1991). If individual experience is the chief determinant of odor connotations, our results suggest a far greater commonality of olfactory experience than is usually acknowledged. Alternatively, in Rader and Tellegen's (1987) theoretical terms, odor-visual correspondences may have, on their own, a normative component, which allows them to overwhelm associations to idiosyncratic personal experiences.

Another explanation for the observed color-odor consensus is that the cross-modal effect is mediated by semantic involvement (Melara, 1989). In the context of Experiment 1, this account would suggest that a smell activates a semantic label, which in turn activates a related color concept. The literature on odor perception suggests certain theoretical and empirical difficulties with this account. First, the tip-of-the-nose phenomenon, in which a subject cannot name an odor despite judging it highly familiar (Lawless & Engen, 1977), has been interpreted to be the result of a fundamental difficulty in retrieving semantic labels for odors. This implies that semantic labels were unlikely to be readily available as an intermediate processing step to color associations. Second, certain of the test odors (e.g., 2–ethyl fenchol and galbanum) have no commonly recognized semantic label, are not frequently encountered by most people, and yet produced significant color score differences.

Third, the smell of lavender, which on a verbal-mediation account should have produced a purple or blue connotation, did not do so.

These results do not rule out a semantic mediation model. Color profiles for certain odors can be interpreted as consistent with the model. For example, cinnamic aldehyde scored high for red, and one can speculate that this was mediated by the familiar term "red hots" used to describe cinnamon-flavored candies. Given these results, it seemed that a nonsemantic approach to color-odor correspondences would be valuable. This was explored in Experiment 2.

The perceptions evoked in individuals with extreme synesthesia show high test-retest reliability (Rizzo & Eslinger, 1989). However, the normative results of Rader and Tellegen (1987) have not previously been examined for consistency over time. Based on a 2–year test-retest interval, we provide the first estimates of consistency for the Rader and Tellegen cross-modal tasks. Mean results at initial and repeat testing were moderately well correlated ($.50 \leq rs \leq .80$). That the correlations were not as high as those found for individual synesthetic subjects (e.g., Rizzo & Eslinger, 1989) is not surprising, given that our test subjects were drawn from the general population, not a clinical one. Based on a procedure like that of Rader and Tellegen, the colors used to describe odors demonstrated a test-retest correlation of $r = .53$. Thus, although color-odor correspondences are as reliable as the line-color correspondences, they are not as high as for the two other nonolfactory tests.

Comparison of test-retest correlation coefficients and task demands may provide some insight as to the nature of the intervening cognitive processes. The highest test-retest correlation was found for the mood-color task, in which both stimuli (mood terms) and descriptors (color names) were presented semantically. In the other three tasks, the stimuli were nonsemantic (line elements or odors) and the correlations were lower. This pattern of results is compatible with the semantic mediation model of cross-modal matching. The intervening associations from stimulus to description were readily accessible between semantic labels for mood and color. In contrast, nonverbal stimuli require the activation of a semantic label prior to additional processing. This could account for the greater test-to-test variability for lines and odors.

EXPERIMENT 2

The task used in Experiment 1 required semantic processing of color names, and the results may have been constrained by the selection of color descriptors. The exclusively semantic basis of the description task also left open the possibility that subjects had different colors in mind when they endorsed, for instance, the term *red.* These considerations led us to design a visual processing task to assign colors to an odor. This task permits direct comparison of color descriptions between subjects, and it has the additional advantage of quantifying color descriptions objectively and precisely. In Experiment 2, the Munsell color specification system was used to match odor stimuli to color samples specified by hue, saturation, and lightness.

METHOD

Subjects

Subjects were 24 men and 26 women (18–40 years, $M = 28 \pm 7$ years). All were recruited and screened as in Experiment 1, and none had participated in that experiment.

Materials

Color samples consisted of all 1,487 glossy finish chips from the *Munsell Book of Color* (Macbeth Division, Kollmorgen Instrument Corp., New Windsor, NY) and 78 chips from the *Supplementary 80–Hue Colors*. The Munsell system describes a three-dimensional color space based on hue, chroma, and value. Hues range from 1 to 100 and are schematically represented around the perimeter of a circle. The circle is radially divided into 10 equal sections, each with a nominal descriptor (red, yellow, green, blue, purple, and their five intermediates, e.g., yellow-red). *Chroma* indicates the degree of color saturation and ranges in psychophysically equal steps from 1 (*least saturated*) to 16 (*most saturated*). *Value* indicates the degree of color lightness and ranges from 1 (*black*) to 10 (*white*) in psychophysically equal intervals. Munsell chroma and value correspond to saturation and gray scale, respectively.

Color chips were mounted on white pasteboard pages with margins coded for identification, and the pages were mounted on circular hinges to avoid positional bias. Color chips were arranged with saturation increasing toward the outer edge of the page and lightness increasing toward the top. Pages were arranged by consecutive hue number.

Twenty odor stimuli were prepared and presented as in Experiment 1.

Procedure

Color chips were presented in a portable light box (Macbeth^{R}/Kollmorgen Instruments Corp.,Newburgh, NY, model JR4–201) with a neutral gray interior (Munsell designation N7/), illuminated with D75 and UV lightbulbs. A black felt backcloth was placed against the rear interior wall of the box. Room lighting was turned off during testing.

Prior to testing, the subject was told in nontechnical terms how the samples were arranged within the book, and was given time to examine all the color stimuli. The subject wore cotton gloves to avoid staining the color chips. On each trial, the experimenter placed an odor stimulus bottle on a bracket at nose level near the subject. Bottles were not handled by the subject. The subject was instructed to sniff from the bottle and find the one color chip that best represented the odor. Location of the selected chip was recorded by the experimenter.

Analysis

To compare color selections across odors, mean coordinates were calculated for each odor in three-dimensional color space. The arithmetic mean was used for the Munsell chroma and value parameters. Because the Munsell hue parameter is a circular distribution, the arithmetic mean is not an appropriate measure. Instead, circular statistics were used to calculate a mean hue vector, which consists of the mean vector angle, ϕ, and the mean vector length, r (Batschelet, 1981, Equations 1.3.8, 1.3.9, and 1.3.10). Functionally, ϕ identifies the center of the hue observation distribution, and r reflects the concentration of observations around the hue circle; r varies between 0 (no concentration) and 1 (maximum concentration) for unimodal circular distributions. Using Rayleigh's test of significance, $r \geq .25$ ($n = 50$) was taken as evidence of a nonrandom ($p < .05$) circular distribution (Batschelet, 1981).

RESULTS

Mean hue vectors for the test odors are plotted in Figure 2. Thirteen odors had significant r values, indicating a substantial normative agreement among subjects when selecting hues to describe these odors. The

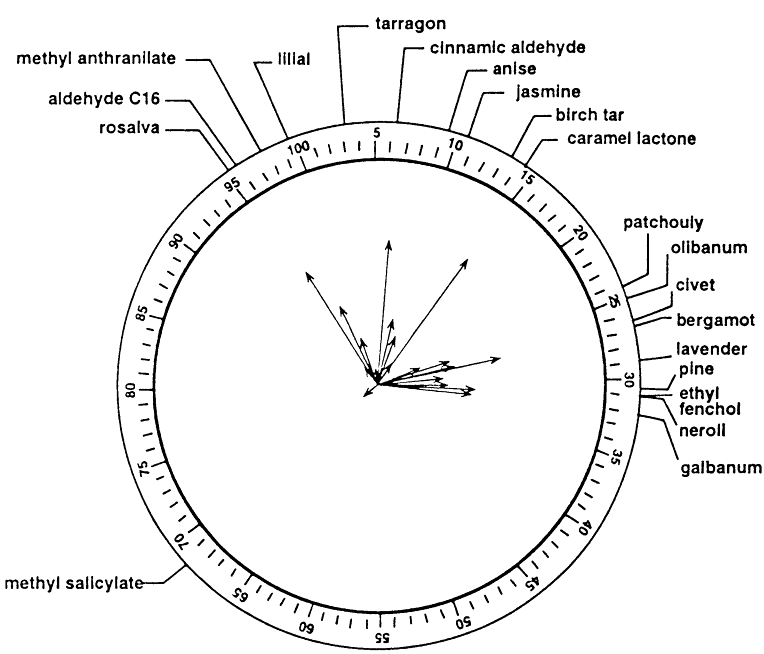

Figure 2. Mean Munsell hue vectors for the 20 odors in Experiment 1 (numbers around the circle represent Munsell hues; radius of inner circle corresponds to vector length of 1.0)

significant hue matches fell within a large range on the hue circle: red-purple, red, yellow-red, yellow, and green-yellow. No color matches were found in the range of green, blue-green, purple-blue, blue, and purple.

There were also differences across odors in the value and chroma of the selected chips. A repeated measures ANOVA using Munsell values was significant, $F(19, 931) = 7.87$, $p < .001$, as was a similar analysis using Munsell chroma, $F(19, 931) = 14.45$, $p < .001$. Mean chromas for the test odors ranged from 3.7 to 10.1 (overall mean 7.0), and mean values ranged from 4.4 to 6.9 (overall mean 5.6). Means for all three Munsell parameters are presented in Table 2.

Using all three mean color parameters, it was possible to locate one Munsell chip that best approximated the group response. The Munsell specifications of the corresponding "best chips" are listed in Table 3 for the 13 odors with significant *r* values.

DISCUSSION

The color-matching task produced significant Munsell hue identifications for 13 of the 20 test odors. The selected hues spanned a rela-

Table 2. Mean Munsell parameters and mean hue vector length (*r*) for the odors tested in Experiment 2

Odor	Hue	Value	Chroma	*r*
Caramel lactone	14.7	5.0	5.1	.68
Cinnamic aldehyde	6.2	4.7	9.3	.66
Aldehyde C-16	96.4	5.5	10.1	.63
Bergamot oil	26.5	6.9	10.1	.54
Neroli oil	31.0	6.5	8.0	.41
Galbanum oil	32.0	4.4	5.2	.40
Methyl anthranilate	97.6	6.0	8.7	.39
Civet artificial	26.4	4.8	3.7	.35
Olibanum oil	25.5	6.2	8.0	.33
2–Ethyl fenchol	30.6	5.3	6.1	.32
Pine oil	30.6	6.0	7.1	.32
Lavender oil	28.8	6.2	7.3	.30
Star anise oil	9.4	4.6	6.2	.29
Jasmine absolute	10.6	6.5	8.2	.24
Lilial	99.2	6.3	8.4	.21
Patchouly oil	24.2	5.0	5.5	.19
Birch tar oil	13.6	5.8	4.4	.11
Rosalva	95.4	6.2	5.7	.10
Methyl salicylate	68.6	6.0	8.0	.08
Tarragon oil	3.1	4.8	5.1	.05

Table 3. Nearest Munsell chip specification for odors with nonrandom hue distributions in Experiment 2, and most frequently endorsed color term in follow-up test

Odor	Hue	Chroma/ value	Modal color name	Exp. 1 rank of color name
Aldehyde C-16	6.25RP	6/12	Pink	1
Bergamot oil	6.25Y	7/12	Yellow	1
Caramel lactone	5YR	5/6	Brown	1
Cinnamic aldehyde	6.25R	5/12	Red	1
Civet artificial	7.5Y	5/4	Green	7
2–Ethyl fenchol	10Y	5/6	Green	1
Galbanum oil	2.5GY	4/6	Green	1
Lavender oil	10Y	6/8	Green	6
Methyl anthranilate	7.5RP	6/8	Pink	1
Neroli oil	10Y	7/8	Green	2
Olibanum oil	5Y	6/8	Green	3
Pine oil	10Y	6/8	Green	1
Star anise oil	10R	5/6	Brown	5

Note. GY = Green-Yellow, R = Red, RP = Red-Purple, Y = Yellow, YR = Yellow-Red.

tively large portion of the hue circle in the range of red-purple through green-yellow. The absence of hue matches to green, blue, and purple may have been due to the particular odors tested. However, it may also reflect the more general case that few odors correspond to these hues. Testing of additional odors is required to address this question.

Because the color-matching task was developed to complement the color term results of Experiment 1, we wanted to know what color terms subjects would use to describe the Munsell color chips selected in Experiment 2, and in particular whether these terms were similar to those generated in Experiment 1. Accordingly, Munsell chips for the 13 odors with nonrandom hue distributions (Table 3) were presented one at a time in a light box against a black felt backdrop under D75 plus UV illumination. Test subjects (19 women and 11 men; *M* age 37 ± 14 years) were asked to mark on an answer sheet 1 of 11 color names that best described the chip. The color names were those used in Experiment 1. Stimulus order was randomized across subjects.

There was close alignment of the color terms generated in Experiment 1, and the color terms used to describe the Munsell chips generated in Experiment 2. For 10 of the 13 chips tested, the most frequently endorsed color name was also a high-ranking color name generated in Experiment 1 (Table 3). Civet and star anise were not associated with their Experiment 1 color names (brown and black, respectively). This may be attributable to the poor representation of brown in the Munsell system, and to the unavailability of pure gray or black chips in Experiment 2. On balance, the converging evidence from Experiments 1 and 2 is consistent with conclusions of Williams and Foley (1968) regarding the interchangeability of color names and color samples. They found that color names and samples produced similar connotative meanings.

In addition to differences in hue, the mean Munsell value and chroma of the selected chips also varied significantly across odors. This implies that subjects were not attending exclusively to the hue dimension; rather, their selection of matching color chips also took into account value and chroma. Thus, our results indicate that subjects may have matched odors to particular shades of color.

GENERAL DISCUSSION

We gave subjects the opportunity to describe an actual odor stimulus in terms of color. The use of a semantic processing task (color names in Experiment 1) and a visual processing task (color chips in Experiment 2) led to similar results. Of the 20 odors tested, all yielded nonrandom color profiles in Experiment 1, and 13 yielded significant hues in Experiment 2. Munsell color specifications and color names were in good agreement. Thus, the two methods provided converging lines of evidence, suggesting that specific odors are associated with specific colors.

How do the color-odor linkages demonstrated here relate to phenomena of sensory equivalence? It would appear that they have more in common with cross-modal correspondences than with synesthesia. It has been held that cross-modal correspondence and synesthesia are separate, though not entirely distinct, phenomena (Marks, 1978). They have also evolved two very different research traditions. The study of cross-modal associations (Marks, 1975), including synesthetic descriptions (Simpson & McKellar, 1955), follows a nomothetic tradition; its central concerns are with how common the perception of sensory equivalence is in the general population, and with how similar the intermodal matchings are among individuals. For example, Rader and Tellegen (1987) reported that certain synesthetic abilities are continuously distributed in the population, and Marks (1975) concluded that associations between color and sound are "often regular, systematic, and consistent from one person to another" (p. 303). It was far from clear that we would be able to find olfactory-based cross-modal correspondences, given the literature on the tip-of-the-nose phenomenon (Cain, 1979; Lawless & Engen, 1977) and the Proustian effect (Ehrlichman & Bastone, 1992). In addition, Marks (1978) had posited that color-odor correspondences would be "erratic and idiosyncratic" (p. 98). The evidence found in this study suggests that color-odor correspondences exist, and that they are on as firm an empirical footing as other forms of cross-modal correspondence, in that they show a similar degree of intersubject consistency.

In contrast to cross-modal associations, the methodological tradition for synesthesia is an idiographic one: it is based on the clinical literature and has stringent defining criteria for cross-modal experience. For example, according to Cytowic and Wood (1982) and Cytowic (1989), an immediate, vividly experienced, involuntary, and relatively invariant percept must be evoked in one modality by stimulation in another. Case studies of individuals with remarkable synesthetic abilities (e.g., Baron-Cohen, Wyke, & Binnie, 1987) highlight key elements of the phenomenon. For example, the color experiences evoked by tones in an individual patient were remarkably stable over 5 months (Rizzo & Eslinger, 1989). How does the reliability of our results compare? Our call-back experiment showed that odor-color correspondences were stable for a group of subjects over a period of 2 years, a consistency of result equal to that observed in a similar, but nonolfactory, task. Few, if any, test-retest studies of normal, nonclinical subjects are available for comparison. However, our results do not approach the degree of invariance seen in clinical cases of synesthesia. This would again seem to place color-odor phenomena under the more general rubric of cross-modal associations.

A question that has been raised, but not definitively settled, in the cross-modal literature is the role of semantic mediation. Melara (1989) has found the semantic account inadequate to explain all the interactions between color and pitch. Color-odor correspondences may be open to a similar interpretation. Some odors may have evoked color connotations via semantic mediation, perhaps with odor recognition as a preliminary step. However, our experiments were not designed to explore the cognitive processing of the response, and the tasks required neither odor recognition nor naming. The role of semantic mediation remains to be determined by additional investigations.

Our demonstration of normative odor-color congruencies establishes an empirical basis for a new set of intermodal relations. Future experiments can address the possibility of relating odor to color on a dimensional basis as has been done for color and sound (Marks, 1975; Melara, 1989). For example, the association of brightness (bright-dim) and loudness (loud-soft) implies that these dimensions are systematically related (Marks, 1987). Although a sensory ordering for odor qualities continues to elude quantification, one can nevertheless imagine relations between color and other sensory dimensions of odor. An example would be color lightness varying systematically with odor intensity. An experiment using the color-matching protocol of Experiment 2 and odors varying in perceived intensity could address this possibility. If dimensional relations can be documented, color-odor correspondences will have a phenomenological description equal in richness to other cross-modal phenomena.

Notes

We thank Harry Methven, Helaine Balsamo, and Joy Drakeford for assistance in carrying out these experiments, Kent Lombard for advice on fragrance materials, Louis Hsu for statistical advice, and John Santelli for reviewing an earlier version of this article.

Correspondence concerning this article should be addressed to Avery N. Gilbert, who is now at Synesthetics, Inc., 113 Inwood Avenue, Upper Montclair, NJ 07043. Received for publication March 17, 1994; revision received October 5, 1994.

References

Baron-Cohen, S., Wyke, M. A., & Binnie, C. (1987). Hearing words and seeing colors: An experimental investigation of a case of synaesthesia. *Perception, 16,* 761–767.

Batschelet, E. (1981). *Circular statistics in biology.* New York: Academic Press.

Burghardt, G. M. (1977). The ontogeny, evolution, and stimulus control of feeding in humans and reptiles. In M. R. Kare & O. Maller (Eds.), *The chemical senses and nutrition* (pp. 253–275). New York: Academic Press.

Cain, W. S. (1979). To know with the nose: Keys to odor identification. *Science, 203,* 467–470.

Cytowic, R. E. (1989). Editorial: Synesthesia and mapping of subjective sensory dimensions. *Neurology, 39,* 349–350.

Cytowic, R. E. (1993). *The man who tasted shapes.* New York: G. P. Putnam's.

Cytowic, R. E., & Wood F. B. (1982). Synesthesia: I. A review of major theories and their brain basis. *Brain and Cognition, 1,* 23–25.

Davis, R. G. (1981). The role of nonolfactory context cues in odor identification. *Perception & Psychophysics, 30,* 83–89.

Ehrlichman, H., & Bastone, L. (1992). Olfaction and emotion. In M. J. Serby & K. L. Chobor (Eds.), *Science of olfaction* (pp. 410–438). New York: Springer-Verlag.

Engen, T. (1972). The effect of expectation on judgments of odor. *Acta Psychologica, 36,* 450–458.

Gilbert, A. N., & Kare, M. R. (1991). A consideration of some psychological and physiological mechanisms of odour perception. In P. M. Müller & D. Lamparsky (Eds.), *Perfumes: Art, science and technology* (pp. 127–149). London: Elsevier Applied Science.

Ishihara, S. (1990). *Ishihara's tests for colour-blindness.* Tokyo: Kanehara.

Lawless, G., & Engen, T. (1977). Associations to odors: Interference, mnemonics, and verbal labeling. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 52–59.

Marks, L. E. (1975). On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin, 82,* 303–331.

Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities.* New York: Academic Press.

Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 384–394.

Martin, L. J. (1909). Über ästhetische Synästhesie [Esthetic synesthesia]. *Zeitschrift für Psychologie, 53,* 1–60.

Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 69–79.

O'Mahony, M. (1986). *Sensory evaluation of food: Statistical methods and procedures.* New York: Marcel Dekker.

Poffenberger, A. T., & Barrows, B. F. (1924). The feeling value of lines. *Journal of Applied Psychology, 8,* 187–205.

Rader, C. M. (1979). *Explorations in synesthesia.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Rader, C. M., & Tellegen, A. (1987). An investigation of synesthesia. *Journal of Personality and Social Psychology, 52,* 981–987.

Rizzo, M., & Eslinger, P. J. (1989). Colored hearing synesthesia: An investigation of neural factors. *Neurology, 39,* 781–784.

Simpson, L., & McKellar, P. (1955). Types of synaesthesia. *Journal of Mental Science, 101,* 141–147.

Williams, J. E., & Foley, J. W., Jr. (1968). Connotative meanings of color names and color hues. *Perceptual and Motor Skills, 26,* 499–502.

Zellner, D. A., Bartoli, A. M., & Eckard, R. (1991). Influence of color on odor identification and liking ratings. *American Journal of Psychology, 104,* 547–561.

Zellner, D. A., & Kautz, M. A. (1990). Color affects perceived odor intensity. *Journal of Experimental Psychology, Human Perception and Performance, 16,* 391–397.