DOI: 10.1111/j.1745-459X.2009.00239.x

AMBIENT LIGHTING MODIFIES THE FLAVOR OF WINE

DANIEL OBERFELD[1,4], HEIKO HECHT[1], ULRICH ALLENDORF[2] and FLORIAN WICKELMAIER[3]

[1]*Department of Psychology*
Johannes Gutenberg-Universität
55099 Mainz, Germany

[2]*Weingut Fritz Allendorf – Georgshof*
Kirchstrasse 69, 65375
Oestrich-Winkel, Germany

[3]*Department of Psychology*
University of Tübingen
Friedrichstraße 21, 72072
Tübingen, Germany

Accepted for Publication February 3, 2009

ABSTRACT

It is well known that the color of a beverage can influence its flavor. We conducted three experiments to investigate the effect of the ambient room color on flavor, while leaving the color of the beverage unaltered. We chose white wine as the beverage and used several methods to fully explore the potential role of ambient light. First, a group of wine buyers made judgments on flavor and global liking while tasting a Riesling on site at a local winery. Ambient color influenced the subjective value of the wine. Wine tasted better in blue or red environments as compared with green and white. A second group was tested in the laboratory. Ambient color modified the taste, but not the odor of the wine. The influence of ambient color on flavor was confirmed in a third experiment using the method of paired comparisons.

PRACTICAL APPLICATIONS

It has become fashionable in architecture and interior design to abandon the classical white or pastel colors that have traditionally been used for wall

[4] Corresponding author. TEL: +49-6131-39-39274; Fax: +49-6131-39-39268; EMAIL: oberfeld@uni-mainz.de

paint. Bright red and orange colors may prevail in a hip hair salon or a bar may have pink illuminated walls. As the private and public places where beverages are consumed change in ambient color, we should expect the taste experience to change as well. Additionally, our data show that the subjective value of the wine and consequently, the amount of money consumers are willing to spend for it can be influenced by ambient color. Thus, the interior design of, for example, a wine bar should take these effects into consideration.

INTRODUCTION

St. Patrick's Day can be quite a sensory challenge to the unexpecting tourist who orders a beer in an Irish pub. Upon ordering a beer, it so happened to one of the authors (H. H.) that the waitress served a beer glass filled with a bright green liquid accompanied by a smile that forbade protest or inquiry. The revolting taste of the beverage was miraculously transformed into normal beer taste upon closing the eyes. The remainder of the beer was consumed with closed eyes. The strong effect that the color of a beverage can have on flavor perception is well known (for a summary see Delwiche 2004). The effect is genuine, it does not depend – at least not entirely – on learned color-flavor associations (see e.g., Zellner and Kautz 1990) but it appears to be limited to certain beverage-color combinations. Also, depending on the situation, the intensity of the perceived odor is monotonically or non-monotonically related to the intensity of the beverage's coloring (DuBose *et al.* 1980; Kemp and Gilbert 1997; Zellner and Whitten 1999). Rather a lot is known about the particular factors that influence the flavor of a beverage as a function of its coloring (for an overview with respect to the taste of wine see Jackson 2002). However, next to nothing is known about the potential influences of colors in the environment upon flavor judgments. In this paper, we pursue the question whether the ambient color of the environment, too, exerts an influence on the flavor of a beverage. We first provide some background information and motivate our choice of wine as the target beverage, and then report three experiments: one conducted on-site in a winery with customers and two conducted in a laboratory environment. All experiments demonstrate an effect of ambient color.

Effects of Beverage Color Go Beyond Simple Flavor Judgments

The color of a beverage seems to be effective at a very basic level of sensory integration. Conscious efforts to ignore the color are largely futile. For instance, Zampini *et al.* (2007) asked observers to discriminate the flavor of solutions that were either colorless, colored appropriately or colored inappropriately. In the latter case, performance was poorer than with colorless or appropriately colored liquids even though the observers were explicitly told to ignore the color.

Beverage color can not only alter the flavor but also influences derived qualities such as the beverage's ability to quench thirst. Liquids consumed from a glass that was tinted blue were judged to be more thirst quenching than liquids in red, green or yellow glasses (Guéguen 2003). Quite surprisingly, sweet drinks were also judged to be better thirst quenchers than sour fruit juices (Clydesdale *et al.* 1992). Effects on expected and experienced refreshment were reported by Zellner and Durlach (2003). The texture of the beverage also seems very important. A sucrose solution is judged to be more than twice as sweet compared with its solidified gelatinous equivalent (Alley and Alley 1998).

If one evaluates the body of literature on the effects of color on odor and taste perception, a complex picture emerges. On the one hand, effects can be found and are at times very strong. There also seems to be a neurophysiological basis for these effects. Österbauer *et al.* (2005) recorded functional magnetic resonance imaging data, while their observers were exposed to odors, colors or to combinations of both inside the brain scanner. Whenever colors and odors were presented together that subjectively constituted a good match in the eyes of the observer, activity in the orbitofrontal cortex and in the insular cortex was observed. On the other hand, not all beverages are equally affected by coloring, and interactions of odor, color and other qualities of the beverage are rather complex. For instance, carbonated water is rather immune to effects of its coloring on taste (Hyman 1983) and strawberry odor has a positive effect on judged sweetness of an aqueous solution of sucrose, but strawberry color does not affect its judged sweetness (Frank *et al.* 1989).

The complexity of the color-odor interaction may be founded in the fact that the process appears to be located at higher cognitive processing stages. For example, it is known that the effects of color on odor assessments of liquids occur rather late during the verbalization phase when naming the odor (Morrot *et al.* 2001, but see also Herz and Engen 1996). A lexical analysis of the attributes that experts use when describing wine revealed that expert wine tasters chose objects of the same color as the wine, rather than attributes that were closer to the olfactory sensations. In keeping with this analysis, the experts used their red-wine vocabulary when describing a white wine that was tinted red by means of an odorless dye.

When it comes to the related but reverse role of verbal preconceptions in the odor perception of beverages, some, but not all color names readily prime certain odors (Koch and Koch 2003). Red and orange are associated with sweet, while green and yellow are associated with sour. In average observers, these associations produce verbal over-shadowing, that is, a color that is verbally provided in a tasting context will influence the odor judgment. Only experts appear to be immune to the effects of naming on the perceived flavor (Parr *et al.* 2002).

Potential Effects of Ambient Color

Surprisingly little is known about how the effects of beverage color come into effect. They could be mediated by an emotional process. For one, the orbitofrontal cortex appears to be involved in the integration of taste and smell with cognitive assessments (Small *et al.* 2004). There is also evidence that the emotional makeup of an observer, both in terms of emotional state as well as in terms of personality, has an influence on odor perception (Chen and Dalton 2005). Neurotic and anxious persons were better able to identify affective odorants, while others excel with neutral odorants. Women responded faster to pleasant odorants than to neutral odorants, although men showed no such effect. On the other hand, when emotions where instilled with movie material, men tended to perceive odors to be more intense when in non-neutral emotional states.

If the effect of color on flavor is mediated or at least facilitated by an induction of a certain emotional state, then the ambient color of the environment should also have an effect on perceived odor. Or put more cautiously, it is important to understand whether effects of ambient lighting on odor perception exist. Knowing their magnitude will contribute to our understanding of how complex odors are formed. To our knowledge, the single empirical study that has addressed this question has asked six enologists and six average adults to perform similarity ratings on three white wines (Sauvageot and Struillou 1997). The wines were presented pair-wise in white, red and green room lighting. Black opaque glasses were used to prevent an effect of ambient color on beverage color. The authors were unable to find an effect of lighting and then went on to pursue effects of wine color on flavor. Such effects were indeed found; a white wine colored red increased the perceived flavor distance between it and another unaltered wine. The authors also found a difference between experts and novices which they interpret to indicate that novices are thrown off by the mismatched color in a more basic manner than are experts.

We chose to follow suit and also use wine as the substance to be judged in our experiments. For one, this would allow us to compare results with those of Sauvageot and Struillou (1997), but more importantly, the status of wine as a sophisticated beverage would guarantee a high level of emotional involvement. Finally, we know more about wine drinking habits and preferences than about preferences toward any other gourmet food or beverage. For instance, a large representative study queried 3,000 German adults as well as several hundred outspoken wine consumers on their preferences and habits (Hoffmann *et al.* 2006). Among others, Hoffmann and colleagues found that despite the common belief that dry Riesling wine is most popular, it is in fact best liked when semi-dry. In a representative poll, almost 50% of German wine consumers reported a preference for semi-dry wine, and about a quarter of the population each preferred dry and sweet Rieslings. Women preferred sweet wine to a larger extent than did men.

To receive a thorough impression of potential effects of ambient color on perceived wine flavor, we conducted a field experiment with clients of a local winery complemented by two laboratory experiments performed on students at Mainz University. Preference judgments and ratings were used in the field experiment and in the first laboratory experiment, and binary choice methods were used in the second laboratory experiment.

EXPERIMENT 1: A FIELD EXPERIMENT OF AMBIENT COLOR AND FLAVOR

Materials and Methods

Subjects. Two hundred and six visitors of the Allendorf winery participated in the experiment voluntarily. Participants who indicated on the questionnaire that they had participated in a wine tasting involving different ambient colors before, had filled out the questionnaire before, were (partially) color blind or did not respond to any of the three former questions were excluded from the data analysis. The remaining participants (75 male, 75 female) ranged in age from 20 to 79 years (mean [M] = 47.5 years, standard deviation [SD] = 13.95 years). Table 1 displays the characteristics of the four different experimental groups (i.e., the subjects tested under the four colors of ambient lighting).

Fisher's exact test indicated that the proportion of men and women did not differ between the four conditions, $P = 0.90$ (two-tailed). Univariate analyses of variance (ANOVA) computed using a Brown–Forsythe testing procedure, which is robust against variance heterogeneity even if group sizes are unequal, showed that the four groups did not differ in terms of their age $F^*(3, 128.85) = 0.141$, $P = 0.935$, their self-ratings concerning their experience with wine (see Procedure) $F^*(3, 128.07) = 0.688$, $P = 0.561$, or their self-ratings concerning their sense of smell and taste, $F^*(3, 141.42) = 1.17$, $P = 0.325$, either.

The participants were naïve with respect to the aim of the experiment.

Materials. The experiment was conducted in a specially designed tasting room at the Allendorf winery in Oestrich-Winkel on the Rhine river.

TABLE 1.
EXPERIMENT 1: CHARACTERISTICS OF THE FOUR EXPERIMENTAL GROUPS

Color	*n*	Gender		Age (years)	Wine experience	Sense of smell/taste
		Male	Female			
Blue	44	23	21	47.98 (13.18)	3.61 (1.88)	4.14 (1.52)
Green	40	18	22	47.13 (15.87)	3.13 (2.10)	4.55 (1.69)
Red	30	16	14	46.33 (15.55)	3.57 (2.18)	4.07 (1.39)
White	36	18	18	48.39 (11.46)	3.72 (1.58)	3.91 (1.60)
Total	150	75	75	47.52 (13.94)	3.50 (1.93)	4.18 (1.56)

The columns *N, male*, and *Female* show numbers of participants. In the remaining columns, means (with standard deviations [SDs] in parentheses) are displayed. The column *wine experience* shows the average self-ratings of the participants' experience with wine collected on the questionnaire (10-point scale, 0 = "very little experienced," 9 = "very experienced"). The column *sense of smell/taste* gives the average self-ratings in response to the question "I take my sense of smell/taste to be 0 = far below average . . . 9 = far above average." The line *total* displays counts, or means and SDs in parentheses across all participants.

TABLE 2.
EXPERIMENT 1: COLORIMETRIC DATA FOR THE FOUR COLORS OF AMBIENT LIGHTING

Color	Luminance (cd/m^2)	CIE 1931	
		x	*y*
Blue	16.8	0.149	0.047
Green	16.4	0.251	0.659
Red	16.0	0.637	0.322
White	16.2	0.315	0.322

CIE, Commission Internationale de l'Éclairage.

The dimensions of the windowless room were 2.75 m × 6.10 m × 4.75 m (width × length × height). One of the smaller sides of the room contained a lighting system producing colored light. On the opposite wall there was a mirror covering the complete wall. The remaining walls and the ceiling were painted white. The room had a wooden floor. The lighting system consisted of white, blue, green and red fluorescent lamps. The luminance of each group of lamps could be controlled individually. The lamps were covered by frosted glass.

Four different colors of ambient lighting were used in the experiment (blue, green, red and white), adjusted for approximately constant luminance. Table 2 shows the luminance and Commission Internationale de l'Éclairage 1931 *xyL* chromaticity values (cf. Wyszecki and Stiles 1982), measured by a Color CAL colorimeter (Cambridge Research Systems, Rochester, Kent, UK) positioned 0.5 m from the illuminated wall.

The wine was a dry Riesling from the Rheingau region (Allendorf Festival 2004, QbA).

The wine was presented in black opaque wine glasses (Sensus, Schott Zwiesel, Germany; volume 299 mL; designed according to DIN 10960) which made it impossible for the participants to judge the color of the beverage. Each glass was filled with 100 mL of wine at a temperature of 6C. The glasses were sealed with a lid.

The questionnaire and the glass containing the wine were provided on a table positioned so that the participant was facing the illuminated wall from a distance of approximately 1.5 m.

Procedure. A between-subjects design was used. Each participant tasted and judged only one glass of wine under one color of ambient lighting. In order to make the experiment manageable for the employees of the Allendorf winery and because the switching of the colors could not be computer-controlled, each color was presented for 2–4 h during 1 day, and each visitor participating during this time span received the same color. The participants were informed that they would taste a beverage and judge it using an anonymous questionnaire. It was emphasized that there were no right or wrong answers but that only the subjective perception and opinion mattered. No time constraint was imposed and the participants were allowed to taste the wine as often as they liked before or while filling out the questionnaire.

The questionnaire consisted of five questions concerning the taste[1] of the beverage, four more general questions concerning the beverage and several questions concerning information about the participant. The questionnaires were provided in English and in German; only the English version is described in the following (the German version is available upon request from the first author).

In order to make the experiment easier for the nonexpert participants, we decided not to differentiate between taste and odor. Consequently, the participants were allowed to both smell and taste the beverage before judging it on the questionnaire so that both orthonasal and retronasal odor perception was involved (Aubry *et al.* 1999). Additionally, we used common terms for the aroma ratings rather than "wine language." The participants were asked to judge the aroma of the beverage in terms of five dimensions: fruitiness, spiciness, bitterness, sourness and sweetness. Ten-point rating scales were used, labeled with the numbers 0–9. For the fruitiness dimension, the end points of the scale were labeled "not fruity" and "very fruity," and the question read "The beverage tastes not fruity (0) . . . very fruity (9)." An analogous format was used for the other dimensions.

[1] We use the term *taste* to comprise the senses of taste and smell.

On two additional 10-point rating scales, the participant judged the *aroma intensity* (end points "very weak [0]" and "very intense [9]") and the *aroma stability* (end points "dissipates immediately [0]" and "lasts a long time [9]").

On another 10-point scale, the participant expressed his or her *global liking* of the beverage ("I like this beverage not at all [0] . . . very much [9]"). Subsequently, the participant specified the *maximum buying price* (MBP), that is, the maximum price he or she would have been willing to pay for a 0.75-L bottle of the beverage, by writing down an amount in Euros.

Finally, the participants specified their age, gender and profession, rated their *experience with wine* ("very little [0]," "very experienced [9]"), and rated their *sense of smell/taste* ("far below average [0]" . . . "far above average [9]").

The participants were asked not to communicate their judgments or the nature of the study to other potential participants.

Results

Effects of Ambient Color on the Ratings. The effect of ambient color on the ratings was analyzed by conducting separate ANOVAs using a Brown–Forsythe testing procedure for each rating scale.

There was a marginally significant effect of ambient color on global liking, $F^*(3, 140.3) = 2.63$, $P = 0.053$. As can be seen in Fig. 1, on average, the participants liked the wine better if it was presented with the illumination set to blue or red rather than to green or white light. For post hoc pair-wise comparisons, two-tailed *t*-tests were conducted using separate standard errors and approximate degrees of freedom because the sample sizes were unequal, and homogeneity of variance could not be not assumed (cf. Moser and Stevens 1992). The tests showed that the global liking ratings obtained in green light were significantly lower than the ratings obtained in red or blue light ($t[62.80] = 2.03$, $P = 0.047$ and $t[78.74] = 2.19$, $P = 0.032$, respectively), and that global liking was marginally significantly lower in white than in blue or red light ($t[76.10] = 1.88$, $P = 0.064$ and $t[61.67] = 1.74$, $P = 0.087$, respectively).

As Fig. 2 shows, a similar pattern of results was found for the effect of ambient color on the maximum price that the participants were willing to pay for the wine. For the analysis, MBPs more than 1.5 times the inter-quartile range beyond the quartiles were taken as outliers (Lovie 1986), resulting in the exclusion of 4 of the total 138 values. An ANOVA showed a significant effect of ambient color on the MBP, $F^*(3, 71.5) = 3.11$, $P = 0.032$. Post hoc pair-wise comparisons indicated that the MBP was significantly higher in blue or red than in green light ($t[69.35] = 2.64$, $P = 0.010$ and $t[31.82] = 2.33$, $P = 0.027$, respectively), and marginally significantly higher in blue or red than in white light ($t[47.96] = 1.69$, $P = 0.098$ and $t[42.85] = 1.93$, $P = 0.060$,

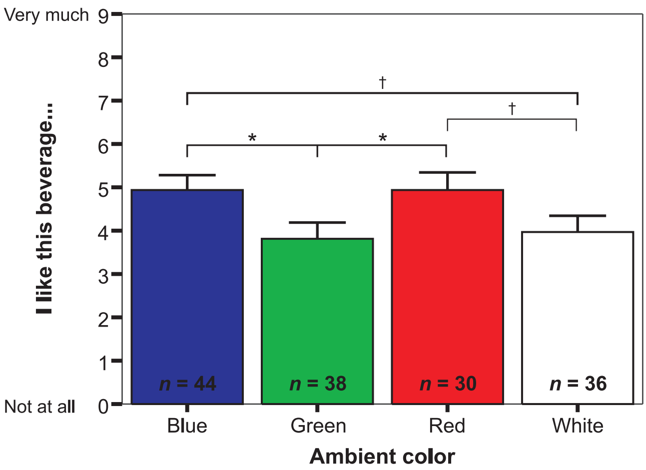

FIG. 1. EXPERIMENT 1: GLOBAL LIKING EXPRESSED ON A 10-POINT RATING SCALE AS A FUNCTION OF AMBIENT COLOR

The label n denotes the number of participants entering the analysis. Error bars show 1 standard error of the mean. Asterisks and crosses indicate significant pair-wise differences according to two-tailed t-tests using separate standard errors and approximate degrees-of-freedom (*: $P < 0.05$, †: $P < 0.1$).

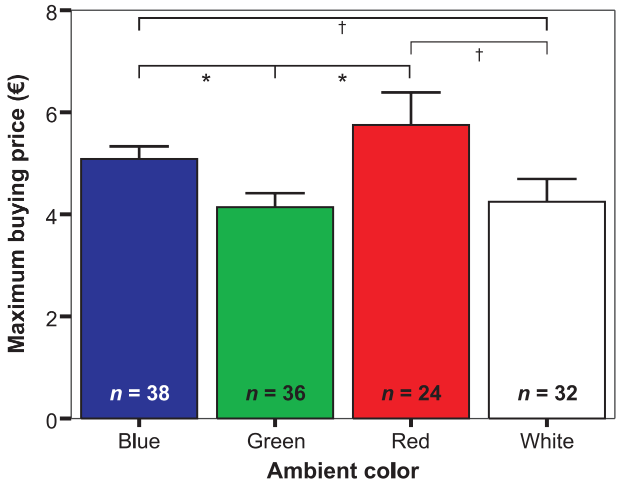

FIG. 2. EXPERIMENT 1: MAXIMUM BUYING PRICE FOR A 0.75-L BOTTLE OF THE WINE AS A FUNCTION OF AMBIENT COLOR. SAME FORMAT AS FIG. 1

respectively). The Pearson product–moment correlation coefficient for the relation between the global liking rating and the MBP was significantly greater than 0, $r = 0.454$, $n = 128$, $P = 0.001$ (two-tailed).

Contrary to the hypotheses, none of the ANOVAs for the aroma ratings (fruitiness, spiciness, bitterness, sourness, sweetness, aroma intensity and aroma stability) showed a significant effect of color of ambient lighting.

Discussion

In sum, we found an effect of ambient lighting on the perceived quality of the wine. Green room light lessened the general liking of the wine and blue or red light made the wine appear more valuable. Interestingly, however, the ambient color did not have any significant influence on the particular aspects of the perceived flavor of the wine.

Two procedural issues could be responsible for the absence of effects on the aroma judgments. First, for nonexperts, it is likely difficult to judge and even more so to rate aroma components of a wine separately, although it should be easier to state whether the wine tastes good or not, that is, to express one's global liking. This should result in a larger variability of the aroma ratings than of the rating of global liking (or the related MBP), making it harder to detect an effect of ambient color. Second, it is well known that our visual system rapidly adapts to changes in ambient lighting, an effect most obvious in the color constancy of the objects in our environment as the spectral composition of the light changes (daylight at noon versus in the evening versus artificial lighting). In other words, it is conceivable that the effect of a change in color would have been more pronounced if the participants had experienced this very change in color. This issue was addressed in Experiment 2, where each participant tasted wines under different colors of ambient lighting.

EXPERIMENT 2: IS PREFERENCE MEDIATED BY EMOTION? A LABORATORY EXPERIMENT

The use of a between-subjects design in Experiment 1 could have resulted in low power for detecting subtle changes induced by ambient lighting. For a given sample size, within-subjects designs usually have more power than between-subjects designs because individual differences in overall rating behavior can be controlled for in the data analysis. For this reason, Experiment 2 utilized a within-subjects design where each participant judged the aroma of the same wine presented repeatedly with differently colored ambient light.

A second methodological aspect of Experiment 1 that may have reduced the effect of color on aroma is that, for reasons that were discussed above, odor and taste were not assessed separately. Koza *et al.* (2005) found that an intense beverage color enhances orthonasal odor intensity (i.e., when the odor is sniffed through the nose). On the other hand, when the beverages were smelled retronasally (i.e., taken in the mouth), color *depressed* odor intensity, compatible with results by Zellner and Durlach (2003). In Experiment 1, these opposing effects may have reduced or camouflaged the effect of a given color on odor. It might have also increased variability due to some subjects basing their ratings predominantly on orthonasal odor and others on retronasal odor. These potential effects were circumvented in Experiment 2 by asking the participants to first assess the (orthonasal) odor of the wine by holding the glass to their nose and sniffing, and subsequently, to judge the flavor (which also includes retronasal odor) by taking a sip of wine into their mouth.

What could be the reason for the effect of ambient color on global liking found in Experiment 1? A straightforward explanation would be that if a color induces a positive mood or emotion (for a discussion of the conceptual distinctions, see Clore *et al.* 1994), that is, if a color makes one feel good, then the same wine tasted in this positive mood is liked better than when in a negative mood. The effects of color on emotion have been studied extensively (Kaiser 1984; Valdez and Mehrabian 1994; Ou *et al.* 2004; Suk 2006), although there is a debate about whether the emotions are evoked directly (i.e., on a physiological basis) by seeing colors or whether they are merely being suggested by way of associations with color (Whitfield and Wiltshire 1990). Our study used a dimensional approach to emotion, according to which emotions can be classified by specific values on several dimensions. This approach dates back to Wundt (1896), who on the basis of philosophical considerations suggested the dimensions Lust-Unlust (appetite-aversion), Erregung-Beruhigung (arousal-sedation) and Spannung-Lösung (pressure-release from pressure). The first empirically founded dimensional concept was presented by Schlosberg (1952), who initially assumed two dimensions (just as Russell 1980), but later three dimensions (Schlosberg 1954). Three dimensions are assumed by most researchers, as for example Osgood *et al.* (1957), who found that the variation in emotional assessments can be accounted for by the dimensions of (or factors) *affective evaluation* (e.g., pleasant versus unpleasant), *activity* (e.g., active versus passive) and *potency* (e.g., strong versus weak). Mehrabian and Russell (1974) suggested the virtually interchangeable fundamental emotional responses: *pleasure*, *arousal* and *dominance*. To test for a relation between the emotional effects of an ambient color and its effect on wine acceptance, pleasure (termed "valence" in the following) and arousal induced by the four ambient colors were assessed using two self-assessment manikin scales (SAM; Lang 1980), both consisting of five comic-like figures. SAM ranges from a smiling, happy figure to a frowning, unhappy figure when representing the valence dimension, and ranges from an excited, wide-eyed figure to a relaxed, sleepy figure for the arousal dimension. Bradley and Lang (1994) showed that the ratings obtained with the SAM scales for valence and arousal are highly correlated to the corresponding ratings (pleasure and arousal) obtained with the Semantic Differential Scale (Mehrabian and Russell 1974).

Materials and Methods

Subjects. One hundred forty-three subjects participated in the experiment voluntarily. Most of them were students at the Johannes Gutenberg-Universität Mainz. Participants who indicated on the questionnaire that they had participated in wine tasting involving different ambient colors before, or were (partially) color blind, were excluded from the data analysis. The remaining 135 participants (69 male, 66 female) ranged in age from 18 to 66 years (M = 26.4 years, SD = 8.6 years). The participants were naïve with respect to the aim of the experiment.

The participants were informed that they would judge several glasses of wine with respect to odor and flavor, and provide their overall impression of each sample. The exact number of wines to be sampled was not specified. Informed written consent according to the Declaration of Helsinki was obtained from all participants prior to the experiment. Specifically, it was emphasized that the consumption of large quantities of alcohol presents a health risk and that it was thus highly recommended not to swallow the wine but to spit it into a container provided inside the box. Participants were also informed that they should not conduct a vehicle in traffic after having participated in the experiment.

On a 10-point rating scale ranging from "I like wine not at all (0)" to "I like wine very much (9)," the average rating was 5.76 (SD = 2.67). The average self-rating of the participants' experience with wine was 3.57 (SD = 2.29), which is not significantly different from the average self-rating of the participants in Experiment 1, $t(263.27) = 0.27$. On the other hand, 43.7% of the participants reported that they drink wine at most once in a month (Fig. 3), and the average buying price for a bottle of wine they normally purchase was only €3.82 (SD = €2.32). Twelve participants indicated that they never buy wine by specifying an average buying price of €0.

Materials. When tasting the same wine several times with only the color of ambient lighting changing, it cannot be precluded that the participants realize that the same wine is presented on each trial, and therefore do not continue to judge the aroma of the wine but simply give the same ratings on each trial. To reduce this problem, two different wines were presented. Both were Riesling wines from the Rheingau region (Allendorf Riesling 2005, Qualitätswein) produced from the same vintage. One of them was of a dry

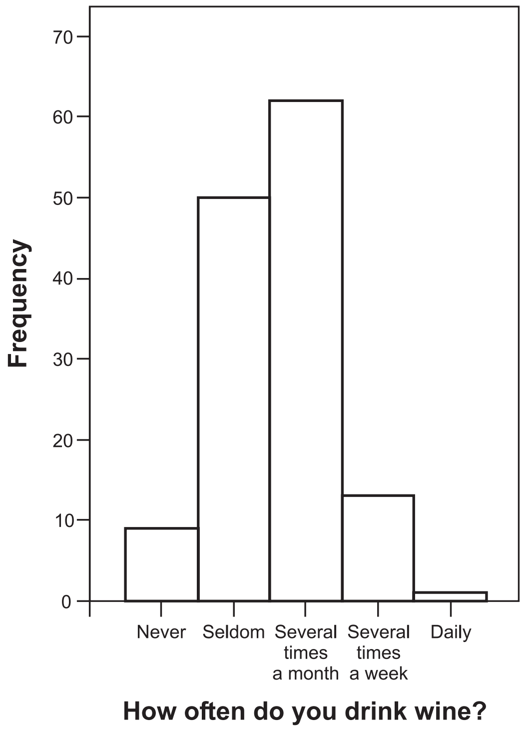

FIG. 3. EXPERIMENT 2: FREQUENCY OF WINE CONSUMPTION

variety, while the other's fermentation was stopped earlier, resulting in a semi-dry wine. Each participant tasted each of the two wines under four ambient colors (blue, green, red and white). The order of the eight Wine × Color combinations was randomized.

Before the participant entered the room, the experimenter prepared eight black opaque wine glasses, each filled with 50 mL of either the dry or the semi-dry wine at a temperature of approximately 10C. For each participant, glasses containing the same type of wine were filled from the same bottle. Each glass was sealed with a lid until the participant received it for the tasting.

The experiment was conducted in a dark room. The participant sat inside a semi-closed cubicle, the front of which was a 98 cm × 98 cm rear projection screen with a projection area of 94 × 94 cm made of frosted glass. The screen was mounted on a table 70 cm high. The ceiling and the two side walls were 120 cm deep and made of white-coated chipboard. The back and the bottom of the cubicle were open. The participant sat on a chair in front of the table holding the projection screen, facing the latter at a distance of approximately

60 cm. The projection screen and a BenQ PB8250 projector (Taipei, Taiwan) connected to a personal computer equipped with an ATI Radeon X500 graphics adaptor (Sunnyvale, CA) were used for the illumination, for the display of instructions, and for the display of the response items. The participant responded via a computer mouse and a computer keyboard positioned in front of the projection screen.

The emotional effect of a color is related not only to its hue, but also to its chromaticity (also termed saturation; Valdez and Mehrabian 1994; Suk 2006). Because of the laboratory conditions of Experiment 2, it was possible to equalize the three chromatic lighting colors in both brightness and chromaticity. The CIE LCH_{ab} color system was used, which is a transformation of the CIELAB system (CIE 1976; cf. Wyszecki and Stiles 1982). The standard illuminant D65 was used as the reference white. In the LCH_{ab} system, each color is represented by the coordinates' lightness (L_{ab}*), chroma (C_{ab}*) and hue angle (h_{ab}*). Lightness denotes the brightness relative to the reference source and ranges from 0 to 100, with the standard illuminant corresponding to $L_{ab}^* = 100$. A chroma value C_{ab}* of 0 means achromaticity (i.e., neutral gray, black or white). The hue angle h_{ab}* signifies hues ranging from red (0°) through yellow (90°), green (180°) and blue (270°) back to red. The characteristics of the four lighting colors are displayed in Table 3. For the measurement, a ColorCAL colorimeter was attached to the projection screen at the approximate eye height of the observers.

The experimenter handed the participant one glass at a time through a small hatch in one wall of the cubicle. The participants could neither see the remaining glasses, nor the bottles from which the glasses were filled.

Procedure. After having read the instructions, which appeared in white ink on the black projection screen, the participant ate a piece of white bread and drank some water to neutralize the taste in his or her mouth. Subsequently, the first block was started. The experimenter switched on the first

TABLE 3.
EXPERIMENT 2: COLORIMETRIC DATA FOR THE FOUR LIGHTING COLORS ACCORDING TO THE *CIELAB* SYSTEM

Color	L (cd/m^2)	a^*	b^*	L_{ab}^*	C_{ab}^*	h_{ab}^* (°)
Blue	43.42	28.9	−95.4	71.8	99.65	286.85
Green	44.07	−65.9	68.6	72.3	95.1	133.85
Red	42.51	78.5	54.9	71.2	95.8	35.0
White	43.53	−1.6	−0.1	71.9	1.6	–

Reference white: D65. L: luminance. L^*: lightness. C^*: chroma. h^*: hue angle.

color and afterwards presented the first glass of wine to the participant. The participant then smelled the wine by holding the glass to his or her nose and sniffing. In order to obtain a more intense odor perception, participants were instructed to swirl the wine in the glass, sniff for one or two seconds and repeat this sequence at least two times. They then judged the odor of the wine on six 10-point rating scales (fruitiness, spiciness, sourness, sweetness, odor intensity and odor pleasantness). The rating scales were constructed in the same way as in Experiment 1 (e.g., "The wine smells not fruity [0] . . . very fruity [9]") and appeared on the projection screen as radio buttons labeled in black ink. The participants selected 1 of the 10 radio buttons by clicking with the left mouse button and then confirmed their choice by clicking the right button. They were allowed to smell the wine again as often as they liked during the rating process.

Following the six odor ratings, the participant was asked to take a sip of wine, to keep the wine in the mouth for some time, and then to spit it into the container. The participants then rated the flavor on seven scales (sweetness, sourness, spiciness, fruitiness, bitterness, flavor intensity and flavor pleasantness). Finally, they rated the temperature of the wine ("cold [0]" to "warm [9]"), rated their global liking of the wine on the same 10-point scale as in Experiment 1 and typed in the maximum buying price, that is, the amount in Euros they would be willing to pay for a 0.75-L bottle of wine.

The projection screen then turned black, and to neutralize the taste of the preceding wine, the participant was instructed to drink a sip of water and had the opportunity to eat a piece of white bread. Then, the next block started. This procedure was repeated until the eight wine type × color combinations had been presented. Subsequently, the participants indicated how many different wines they thought they had just tasted by selecting a number between 1 and 8 from a drop-down menu.

In the next step, the participants judged the colors. The first color was switched on and the participant gave his or her liking on a rating scale ("I like this color not at all [0] . . . very much [9]"). Next, valence and arousal were assessed using two SAM scales (Lang 1980), which appeared on the screen in black ink, together with the sentence "With this color, I feel . . ." The participants were informed that manikins representing different emotional states ranging from happy to unhappy and aroused to relaxed, respectively, would be presented. The participants chose the scale value corresponding to their momentary emotional state by selecting one of nine capital letters (A–I) presented below the manikins. Five of the letters (A, C, E, G and I) were located vertically centered below the five manikins, and the remaining four letters were centered horizontally between an adjacent pair of the former letters. To identify the anchors of each dimension, the end points of the scales were labeled with representative adjectives. To give an example, the positive

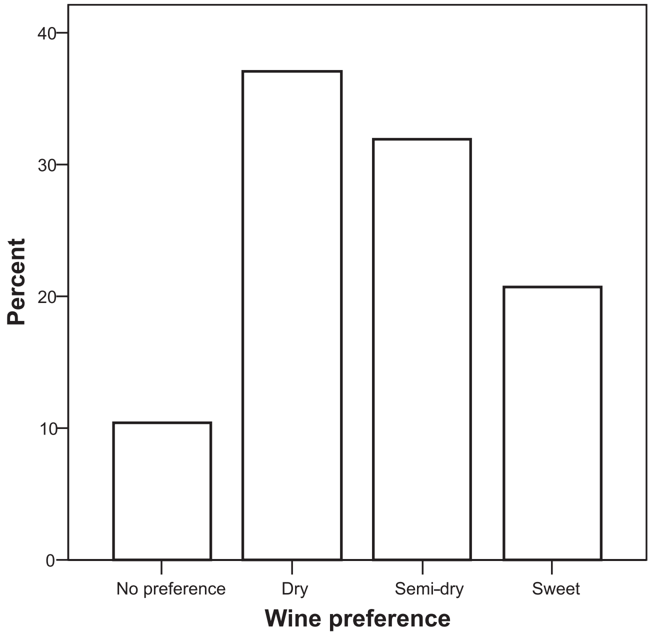

FIG. 4. EXPERIMENT 2: PREFERENCE FOR WINE TYPE

end point of the valence scale was labeled "glücklich (happy), angenehm (pleasant), fröhlich (joyous), vergnügt (cheerful)" (Bradley and Lang 1994). This sequence was repeated for each of the three remaining colors. The order of colors was randomized.

Finally, the participants specified their age, gender and profession, and answered questions regarding their wine consumption and wine preferences (see Section Subjects) (Fig. 4).

Results

Sensory Ratings of the Wine. The effects of ambient color and wine on the ratings were assessed by 2 (wine) $\times$ 4 (color) repeated-measures of ANOVAs conducted separately for each attribute. A univariate approach with the Huynh–Feldt correction for the degrees of freedom was used (cf. Keselman *et al.* 2001). The significant effects are listed in Table 4. We first discuss the effects of wine, and then the effects of color.

The odor of the semi-dry wine ($M = 5.15$, $SD = 1.41$) was rated as being significantly more intense than the odor of the dry wine ($M = 4.85$, $SD = 1.42$). Significant effects of wine were also observed for all flavor ratings. As can be

TABLE 4.
EXPERIMENT 2: RESULTS OF REPEATED-MEASURES ANALYSIS OF VARIANCE

Attribute		Effect	
		Wine (1 and 134 *df*s)	Color (3 and 402 *df*s)
Odor	Intensity	$F = 6.46$, $P = 0.012$	
Flavor	Sweetness	$F = 41.31$, $P = 0.001$	
	Sourness	$F = 26.60$, $P = 0.001$	
	Spiciness	$F = 11.12$, $P = 0.001$	$F = 3.14$, $P = 0.026$, $\tilde{\varepsilon} = 0.99$
	Fruitiness	$F = 33.14$, $P = 0.001$	$F = 2.39$, $P = 0.071$, $\tilde{\varepsilon} = 0.96$
	Bitterness	$F = 13.70$, $P = 0.001$	$F = 2.32$, $P = 0.075$, $\tilde{\varepsilon} = 1.0$
	Intensity	$F = 21.78$, $P = 0.001$	
	Pleasantness	$F = 13.50$, $P = 0.001$	
Temperature		$F = 4.28$, $P = 0.040$	
Global liking		$F = 11.03$, $P = 0.040$	
MBP		$F(1, 125) = 15.81$, $P = 0.001$	

Only effects with $P < 0.1$ are listed. Note the deviating *df*s for the attribute maximum buying price (MBP), due to the exclusion of outliers (see text).

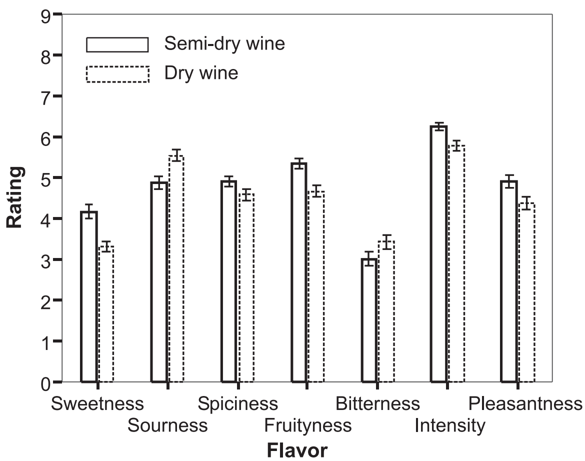

FIG. 5. EXPERIMENT 2: MEAN RATINGS ON THE SEVEN FLAVOR ATTRIBUTES, AS A FUNCTION OF WINE (SOLID LINES: SEMI-DRY, DASHED LINES: DRY)
The differences between the two wines were significant ($P < 0.05$, two-tailed) for all attributes. Error bars show ± 1 standard error of the mean.

seen in Fig. 5, the semi-dry wine was rated as being sweeter and less sour than the dry wine. The semi-dry wine was also rated as being more spicy and fruity, and less bitter. Its flavor was perceived as being more intense and more acceptable. Given these results, it is not surprising that global liking was

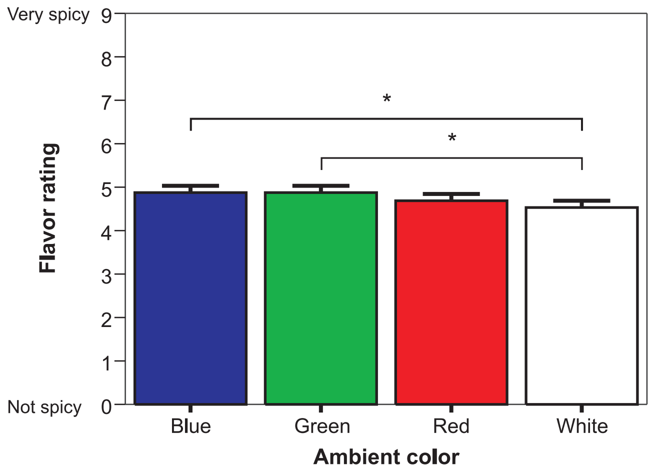

FIG. 6. EXPERIMENT 2: MEAN FLAVOR RATINGS ON THE ATTRIBUTE SPICINESS AS A FUNCTION OF AMBIENT COLOR

Error bars show ± 1 standard error of the mean. Asterisks indicate significant pair-wise differences (two-tailed paired-samples *t*-tests; same format as Fig. 1).

significantly higher for the semi-dry wine ($M = 4.62$, $SD = 1.87$) than for the dry wine ($M = 4.11$, $SD = 1.85$). This finding is compatible with the fact that the majority of participants responded that they generally prefer semi-dry or sweet rather than dry wine (Fig. 4). For the analysis of the maximum buying price specified for a 0.75-L bottle, MBPs more than 1.5 times the inter-quartile range beyond the quartiles were taken as outliers (Lovie 1986), thus resulting in the exclusion of 26 of the total of 1,080 data points. As a consequence, the data from only 126 of the total of 135 participants entered the two-factorial ANOVA. The MBP was significantly higher for the semi-dry than for the dry wine ($M = €3.18$, $SD = €1.53$ and $M = €3.54$, $SD = €1.59$, respectively). Finally, on a scale ranging from "cold (0)" to "warm (9)," the semi-dry wine was rated to be slightly but significantly warmer than the dry wine (M = 3.05, $SD = 1.22$ and $M = 2.91$, $SD = 1.31$, respectively).

How about the effect of color on the ratings? Contrary to expectation, ambient color had no significant effect on any of the odor ratings. For the flavor ratings, however, there was a significant effect of ambient color on spiciness (Table 4). As Fig. 6 shows, the participants perceived the wine to be spicier if it was presented in blue or green rather than in red or white light. Post hoc pair-wise comparisons were conducted using paired-samples *t*-tests. In both the blue and the green light, the spiciness ratings were significantly higher than in the white light ($t[134] = 2.45$, $P = 0.016$ and $t[134] = 2.53$, $P = 0.013$, respectively).

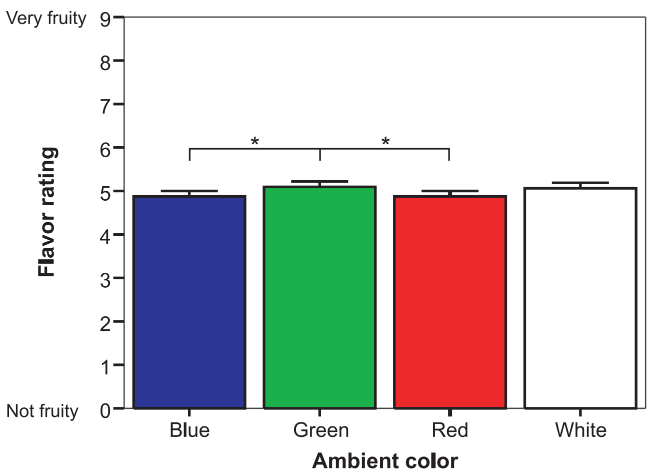

FIG. 7. EXPERIMENT 2: MEAN FLAVOR RATINGS ON THE ATTRIBUTE FRUITINESS AS A FUNCTION OF AMBIENT COLOR. SAME FORMAT AS FIG. 1

Ambient color also had a marginally significant effect on fruitiness, with the ratings in green and white light being higher than in blue and red light (Fig. 7). Post hoc pair-wise comparisons showed that the rating obtained in blue and red differed significantly from the ratings obtained in green light ($t[134] = 2.23$, $P = 0.027$ and $t[134] = 1.99$, $P = 0.049$, respectively). Essentially, the opposite pattern was observed for bitterness (Fig. 8), where the flavor was rated to be more bitter when the wine was presented in blue rather than in green or white light ($t[134] = 2.55$, $P = 0.012$ and $t[134] = 1.80$, $P = 0.075$, respectively).

Surprisingly, unlike in Experiment 1, there was no significant effect of ambient color on global liking, $F(3, 402) = 0.56$. The MBP showed no influence of ambient color either $F(3, 375) = 0.83$.

The Wine × Color interaction was nonsignificant for all dependent variables.

Because the above univariate analyses showed effects of color only on the flavor ratings, but none on the odor-ratings, a doubly multivariate repeated measures ANOVA was used to test for effects of ambient color, wine and their interaction on the combination of the seven flavor ratings. Ambient color had a marginally significant effect, Pillai's trace = 0.229, $F(21, 114) = 1.61$, $P = 0.059$. The effect of wine was significant, Pillai's trace = 0.372, $F(7, 128) = 10.85$, $P = 0.001$. The Wine × Color interaction did not reach significance, Pillai's trace = 0.140, $F(21, 114) = 0.880$.

Further evidence for an effect of ambient color comes from the distribution of responses to the question of how many different wines the

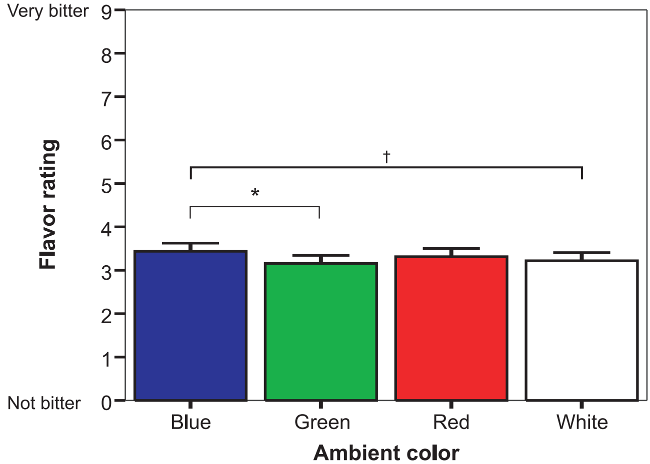

FIG. 8. EXPERIMENT 2: MEAN FLAVOR RATINGS ON THE ATTRIBUTE BITTERNESS AS A FUNCTION OF AMBIENT COLOR. SAME FORMAT AS FIG. 1

participants thought they had sampled during the experiment. As Fig. 9 shows, only 19.3% of the subjects correctly stated that they had tasted two different wines. Fourteen percent thought to have tasted only one wine, although the remaining 58.5% of the participants had the impression of having been presented more than two wines. A one-sample t-test indicated that the mean estimated number of wines was significantly greater than 2, $t(134) = 9.29$, $P = 0.001$ (one-tailed).

Emotional Effects of the Colors. A repeated-measures ANOVA showed that the liking ratings obtained for the four ambient colors differed significantly, $F(1, 134) = 80.40$, $P = 0.001$, $\tilde{\varepsilon} = 0.978$. As the triangles in Fig. 10 show, on average, the participants liked the blue illumination the best, followed by red, green and white, a finding which is compatible with results on color preferences in western adults (for recent studies see Terwogt and Hoeksma 1995; Crozier 1999; Dittmar 2001; Camgöz *et al.* 2002). Post hoc pair-wise comparisons (paired-samples t-tests) indicated significant differences ($P < 0.001$, two-tailed) between all pairs of colors except green and red.

What can be concluded about the color-induced emotions assessed by the SAM scales for valence and arousal? For the purpose of analysis, the nine scale values (A–I) were converted to equally spaced integer values ranging from 1 to 9. For the valence scale, the positive end point was assigned the number 9, and the negative end point was assigned the number 1. For the arousal scale, the end point representing maximum arousal was assigned

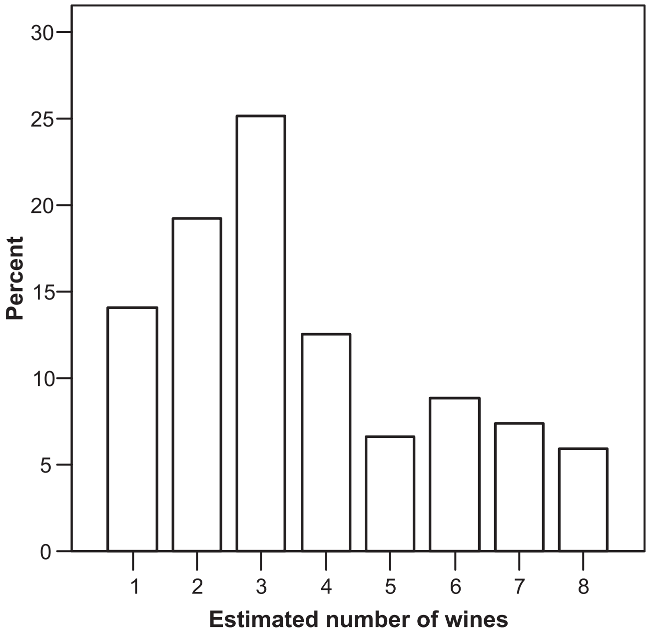

FIG. 9. EXPERIMENT 2: DISTRIBUTION OF THE ESTIMATED NUMBER OF DIFFERENT WINES TASTED DURING THE EXPERIMENT

the number 9, and the end point representing minimum arousal was assigned the number 1. A separate repeated-measures ANOVA was conducted for each scale to analyze the effect of ambient color on the ratings. Mean data are displayed in Fig. 10. Ambient color had a significant effect on valence (boxes in Fig. 10), $F(1, 134) = 47.35$, $P = 0.001$, $\tilde{\varepsilon} = 0.971$. Post hoc pair-wise comparisons indicated significant differences ($P < 0.05$, two-tailed) between all pairs of colors except green and red. Not surprisingly, the pattern of results was similar to the liking ratings (triangles in Fig. 10). To analyze the correlation between the liking rating and valence, the "within-subjects" correlation coefficient was computed by removing the differences between subjects (Bland and Altman 1995), because calculating the correlation coefficient as if the data were a simple sample (i.e., comprised of $n \cdot k$ independent observations, where n is the number of subjects, and k is the number of colors judged by each subject) can be misleading (cf. Bland and Altman 1994). Technically speaking, an analysis of covariance was conducted, with valence as the dependent variable, the liking rating as a continuous covariate and subject as a fixed factor. The magnitude of the within-subjects correlation coefficient is then calculated as $\sqrt{SS_{\mathrm{Liking}} / (SS_{\mathrm{Liking}} + SS_{\mathrm{Error}})}$ (Bland and Altman 1995), and the

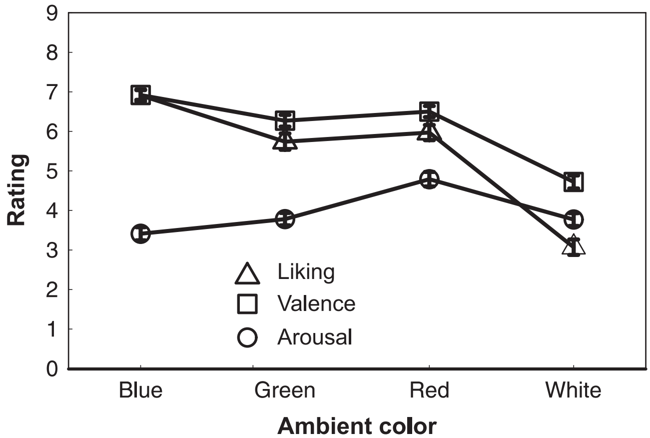

FIG. 10. EXPERIMENT 2: MEAN RATINGS OF THE LIKING AND THE EMOTIONAL EFFECT OF THE FOUR LIGHTING COLORS

Triangles: Liking (0 = min., 9 = max.). Boxes: self-assessment manikin (SAM) scale valence (1 = negative, 9 = positive). Circles: SAM scale arousal (1 = min., 9 = max.). Error bars show ± 1 standard error of the mean.

sign of the correlation coefficient is given by the sign of the regression coefficient for liking. There was a significant positive correlation between the liking rating and valence, $r = 0.73$, $F(1, 404) = 427.5$, $P = 0.001$.

Arousal was also influenced significantly by ambient color (circles in Fig. 10), $F(1, 134) = 15.10$, $P = 0.001$, $\tilde{\varepsilon} = 1.0$. Post hoc pair-wise comparisons showed that the red illumination caused a significantly higher level of arousal than blue light ($t[134] = 6.17$, $P = 0.001$) or green light ($t[134] = 4.87$, $P = 0.001$). The peak in arousal associated with red light is consistent with the effects of color on physiological parameters like the galvanic skin response or systolic blood pressure (Kaiser 1984), which are related to arousal and stress/relaxation.

There was a weak negative (within-subjects) correlation between valence and arousal, $r = -0.14$, $F(1, 404) = 7.5$, $P = 0.006$.

The emotional effects of the colors were assessed to test the hypothesis that the overall liking of the wine is related to the emotional effects of the color. To answer this question, a multiple regression analysis was conducted. The two predictors were valence and arousal, and the dependent variable was the global liking rating. As the data are from a repeated-measures design where each subject contributed eight (Wine $\times$ Color) global liking ratings, it is likely that these observations are correlated, and thus the ordinary-least-squares multiple regression analysis assuming independent observations is inappropriate (see Burton *et al.* 1998). For this reason, generalized estimating equations

(GEE; Liang and Zeger 1986) were used. In this approach, the marginal (or population-averaged) expectation of the dependent variable is modeled as a function of the predictors (covariates). The covariates are related to the marginal probabilities, treating the structure of the correlations between observations obtained from the same subject as a nuisance parameter. Note that an alternative approach are subject-specific models which allow parameters (e.g., the intercept and the slope in regression) to vary from subject to subject. An example for the latter type of models are random-effects models, which model the correlation structure by treating the subjects as a random sample from a population of all such subjects (for a discussion of marginal models compared with subject-specific models see Liang and Zeger 1993; Pendergast *et al.* 1996). An advantage of the GEE method is that it uses a "robust" (or "empirically-corrected") estimate of the variance-covariance matrix that is consistent even under misspecification of the "working" correlation matrix used when fitting the model (Overall and Tonidandel 2004). Therefore, GEEs offer the potential of providing asymptotically unbiased estimates of the regression parameters, even in cases where the exact nature of the intra-subject dependence is unknown (Zorn 2001). The procedure GENMOD of the statistics software SAS (SAS Institute, Cary, NC) was used for the analysis. The working correlation matrix was specified as being of type "unstructured," that is, the procedure placed no constraints on the correlations across observations within a subject. The two independent variables (predictors) were entered simultaneously.

Neither the regression coefficient for valence nor the regression coefficient for arousal differed significantly from 0 ($\beta_{\text{Valence}} = 0.060$, empirical $SE = 0.037$, $z = 1.62$, $P = 0.106$ and $\beta_{\text{Arousal}} = 0.006$, empirical $SE = 0.033$, $z = 0.18$, $P = 0.86$, respectively). Thus, it can be concluded that global liking of the wine is not related in a simple, linear manner to the emotional effects of the color.

Discussion

By asking observers to first assess the smell of the wine and then to make flavor judgments, we sought to pinpoint the influence of ambient color. As the distinction between odor and flavor may be rather subtle to the untrained observer, we also resorted to a laboratory setup and a repeated-measures design, which should be more powerful in detecting even small effects. Contrary to the positive effects of beverage color on orthonasal odor perception (Koza *et al.* 2005), ambient color had no effect on odor. The retronasal flavor judgment, however, was significantly modified by ambient color. Blue and green room lighting made the wine taste spicier and somewhat fruitier. Interestingly, blue lighting made the wine taste bitter but at the same time the blue light resulted in larger liking scores of the wine. This apparent contradiction can be resolved if one assumes that the liking of the color (blue is the predominant favorite color) is confounded with the liking of the wine.

We can only speculate why these effects do not surface when observers merely smell the wine. Maybe when concentrating on the odor judgment, the visual modality is suppressed although it is active when it comes to the integral flavor experience that includes smell, taste, and tactile cues.

Unlike in Experiment 1, there were also no significant effects of ambient color on the MBP. The different characteristics of the sample was the likely cause, as the participants in Experiment 2 were mainly young students who reported to consume wine only rarely and to spend only low amounts of money on wine. Another difference between Experiment 1 and Experiment 2 was the immersiveness of the ambient lighting. The illuminated wall at the Allendorf winery was much larger, and more uniformly illuminated than the translucent screen used in the laboratory setup. Additionally, the mirrors in the color room at the Allendorf winery enhanced the impression. It is unlikely however that this difference in immersion can explain the lack of effects on global liking and MSP as color had a clear effect on flavor in the laboratory setup.

Taken together, these results certainly suggest that the nonexpert participants had problems rating the odor, which resulted in a large variability of the ratings and could have masked potential effects. Flavor judgments, on the other hand, were more precise and did reveal an effect of ambient color.

EXPERIMENT 3: A MORE SENSITIVE LABORATORY METHOD

Experiments 1 and 2 have demonstrated effects of ambient color on the perceived aroma of white wine. The effects appeared to be less stable than effects of wine color and they relied on observers actually tasting the wine as opposed to merely sniffing it. The aim of Experiment 3 was to use a more sensitive method that is able to detect small color-induced flavor differences. The within-subjects design of Experiment 2 resulted in an increased power to detect color-induced aroma changes, however, the rating procedure used in Experiments 1 and 2 may not have been optimal for detecting small changes in aroma. For this reason, a much simpler task was used in Experiment 3. Each participant tasted only two glasses of wine, each of them under a different ambient color, and performed a simple paired comparison. For example, he or she decided whether the first wine or the second wine had tasted better. Such a qualitative, forced-choice decision should be easier to make than a graded response, as for example whether a given wine deserves a “4” or a “5” for fruitiness.

The data were then analyzed using the Bradley–Terry–Luce (BTL) model (Bradley and Terry 1952; Luce 1959). This model aims at predicting the probability that the wine under a certain color of lighting is chosen with respect to a given attribute. To that effect, a latent psychological variable is defined (e.g., fruitiness), and the values on this variable are estimated when the model is fit to the data. Given an adequate fit of the model, the estimated values serve as ratio-scale values on the fruitiness scale. Furthermore, by employing an extension of the BTL model (Davidson and Beaver 1977), it was possible to quantify the effect of the presentation order within a pair.

Materials and Methods

Subjects. The experiment was conducted in a room located near the entrance to the refectory of the Universität Mainz. The subjects were recruited by asking them whether they were interested in participating in a special wine tasting conducted by the Department of Psychology. Informed written consent according to the Declaration of Helsinki was obtained from all participants prior to the experiment. Originally, the experiment was targeted at a sample size of 35 subjects in each of the six experimental conditions (Color Pair × Order, see below). After the data from 215 subjects had been obtained, inspection of the data showed that the strongest differences in preference were present for the color pair blue-red. Therefore, to obtain more reliable estimates for this color pair, 20 additional subjects were tested in the two conditions involving the latter color pair.

Participants who indicated on the questionnaire that they had participated in this, or a similar experiment before, were (partially) color blind or did not respond to one of the two former questions were excluded from the data analysis. The data of one participant who indicated on the questionnaire to be unable to decide between the two samples of wine were also excluded. The 230 remaining participants (84 male) ranged in age from 18 to 59 years ($M = 24.1$ years, $SD = 6.0$ years). The proportion of men and women did not differ between the six different experimental groups (i.e., the subjects tested under each of the 3×2 sequences of colors), $P = 0.294$ (Fisher's exact test, two-tailed). Univariate ANOVAs showed that the six groups did not differ in terms of their age, $F(5, 224) = 0.636$, $P = 0.672$, their self-ratings concerning their sense of smell and taste (see section Procedure), $F(5, 224) = 1.30$, $P = 0.265$ or their self-ratings concerning their general liking of wine, $F(5, 224) = 1.37$, $P = 0.237$ either. The participants also indicated the frequency of their wine consumption, using the same five categories as in Experiment 2. A proportional odds logistic regression model (McCullagh 1980) showed no significant difference between groups, $\chi^2_5 = 6.75$, $P = 0.24$. Finally, the preference for either red or white wine did not differ significantly between the six groups (Fisher's exact test, $P = 0.167$, $n = 225$, two-tailed).

Materials. The same rear projection screen and semi-closed cubicle as in Experiment 2 was used. Because the room in which Experiment 3 was conducted could not be darkened completely, a black light-blocking curtain was mounted on the back of the cubicle, and the space between the projector (Dell MP3200) and the projection screen was isolated so that no ambient light fell onto the projection screen. A PC notebook was used for the presentation of the colors. During the tasting, the participant sat on a chair in front of the table on which the projection screen was mounted, facing the latter at a distance of approximately 60 cm.

Immediately before the participant entered the cubicle, the experimenter prepared two opaque black wine glasses, each filled with 50 mL of wine from the same bottle. The wine was a dry Rheingau Riesling (Allendorf Festival 2006). The glasses were labeled “1” and “2.”

There were three colors of ambient lighting. For the blue, the red, and the white light, the CIE 1931 *xyL* chromaticity values measured by a ColorCAL colorimeter attached to the projection screen at the approximate eye height of the observers were 0.156, 0.081, 24.59 cd/m^2; 0.584, 0.349, 26.63 cd/m^2; and 0.283, 0.355, 25.92 cd/m^2, respectively.

Procedure. Each participant was randomly assigned to one of the six possible sequences of two colors. For example, if assigned to the color sequence blue-red, the participant tasted the first glass of wine under the blue lighting color, and the second glass under the red lighting color. At the beginning, and after having received the instructions, the participant was asked to eat a piece of white bread and to drink a sip of water to neutralize the taste in her mouth. The participant then entered the cubicle in which the first color had already been switched on. The experimenter handed the first glass through a hatch in the cubicle. The participants were instructed to first evaluate the odor of the wine by swirling the wine in the glass, sniffing for one or two seconds, and repeating this sequence at least two times. Subsequently, they were asked to take a sip of the wine and to keep it in their mouth for some time. In order to equalize the tasting time for the two wines, the projection screen turned black 45 s after the presentation of the first glass, and the experimenter removed the glass. The participant was asked to drink a sip of water in order to neutralize the taste. Then the experimenter switched on the second color and handed the subject the second glass, who tasted the wine in the same manner as before, again with a time constraint of 45 s.

Subsequently, the participant left the cubicle and filled in the questionnaire. Participants first responded which of the two wines they preferred on a four-point rating scale consisting of the alternatives “Clearly wine 1,” “Slightly wine 1,” “Slightly wine 2” and “Clearly wine 2.” On three analogously constructed scales, they also specified which of the two wines they had perceived as fruitier, spicier and sweeter.

Finally, the questionnaire contained several general questions concerning for example age, gender and wine preferences (see section Subjects).

Analysis of the Choice Frequencies. Overall, 21% (193/920) of all ratings fell into the categories "clearly wine 1" or "clearly wine 2." It was expected that the individual locations of the category boundaries between "clearly" and "slightly" are highly affected by judgmental rather than sensory factors (Lukas 1991). Therefore, in order to prevent the analysis from being overly sensitive to such presumably volatile effects, the ratings were dichotomized into binary judgments either favoring the first wine ("clearly/slightly wine 1") or the second wine ("clearly/slightly wine 2"). For each within-pair order, the individual judgments were aggregated across subjects to yield choice frequencies (Table 5).

The choice frequencies were analyzed using an extension of the BTL model (Bradley and Terry 1952; Luce 1959) that accounts for the presentation

TABLE 5.
EXPERIMENT 3: SUMMARY OF THE DICHOTOMIZED CHOICE FREQUENCIES FOR EACH ATTRIBUTE, AGGREGATED ACROSS PARTICIPANTS

Attribute	First stimulus	Second stimulus		
		Blue	Red	White
Preference				
	Blue	–	(27, 18)	(23, 12)
	Red	(32, 13)	–	(24, 11)
	White	(22, 13)	(25, 10)	–
Fruitiness				
	Blue	–	(26, 19)	(22, 13)
	Red	(35, 10)	–	(28, 7)
	White	(21, 14)	(26, 9)	–
Spiciness				
	Blue	–	(24, 21)	(20, 15)
	Red	(21, 24)	–	(17, 18)
	White	(16, 19)	(18, 17)	–
Sweetness				
	Blue	–	(22, 23)	(21, 14)
	Red	(32, 13)	–	(23, 12)
	White	(18, 17)	(19, 16)	–

The first and second entries in the brackets denote how often the first and the second stimulus, respectively, was chosen.

order by a multiplicative order effect (Davidson and Beaver 1977). The BTL model assumes that each stimulus is assigned a scale value u, which represents its perceived strength, such that the probability p_{xy} of choosing stimulus x over y in a paired comparison becomes

$$p_{xy} = \frac{u(x)}{u(x) + u(y)}. \tag{1}$$

In order to quantify the effect of the within-pair presentation order, the BTL model has to be extended. Let $p_{xy|x}$ denote the probability of choosing x over y, given that x has been presented first, then the BTL model with order effect predicts that

$$p_{xy|x} = \frac{u(x)}{u(x) + \vartheta \cdot u(y)} \quad \text{and} \quad p_{xy|y} = \frac{\vartheta \cdot u(x)}{\vartheta \cdot u(x) + u(y)}, \tag{2}$$

where ϑ represents the order effect. If the order effect is less than one, then there is a bias favoring the first presentation interval. If it is greater than one, then there is a bias favoring the second presentation interval.

When employing the BTL model in an analysis, it is of interest whether: (1) the model fits the data adequately; (2) the wines presented under the different color of lighting are perceived to be equal with respect to a given attribute (e.g., fruitiness); and (3) there is an order effect for the first or the second presentation interval. These questions were addressed by likelihood ratio tests (for details see Davidson and Beaver 1977). Data analysis and model fitting was performed via a maximum likelihood approach, using software described in Wickelmaier and Schmid (2004).

Results

Choice Frequencies. Table 5 shows the (dichotomized) choice frequencies for each attribute. To give an example, when the first wine was presented in blue light, and the second wine in red light, then 26 subjects chose the wine presented in blue light as being fruitier, and 19 subjects chose the wine presented in red light. When the presentation order was reversed, only 10 participants chose the wine in blue and 35 chose the wine in red.

This example shows two things: first, there seems to be an effect of the ambient color. In the first pair, the proportion that the wine presented in blue light is chosen is 58% (26/45). In the second pair, the proportion that the wine in red light is chosen is 78% (35/45). Thus, in total, wine presented in red light tasted fruitier than wine presented in blue light. If there was no color effect, those two proportions would be equal. Second, the frequencies indicate that

TABLE 6.
EXPERIMENT 3: LIKELIHOOD RATIO TESTS OF THE GOODNESS OF FIT OF THE BTL MODEL (Eq. 2), OF THE EFFECT OF AMBIENT COLOR AND OF THE ORDER EFFECT

Attribute	Goodness of fit		Effect of color		Order effect (ϑ)		
	$G^2(3)$	P	$G^2(2)$	P	Estimate	$G^2(1)$	P
Preference	1.37	.713	0.58	0.750	0.50	25.65	<0.001
Fruitiness	5.04	.169	3.68	0.159	0.45	33.47	<0.001
Spiciness	0.19	.980	1.23	0.542	0.98	0.02	0.895
Sweetness	1.41	.703	5.11	0.078	0.70	7.15	0.007

there is a pronounced bias towards selecting the wine that is tasted first. Whether or not such effects are significant has to be assessed via statistical tests (see next section).

Scale Values. The goodness of fit of the BTL model with order effect was evaluated using a likelihood ratio test. In this test, the likelihood of the BTL model is compared with the likelihood of a saturated binomial model. A nonsignificant test indicates that the BTL model cannot be rejected. As is shown in Table 6, the BTL model with order effect accounted well for the choice data for each of the four attributes.

Figure 11 shows the estimated BTL scale values for each of the four attributes. The overall color effect was again evaluated by likelihood ratio tests (Table 6). When testing the BTL model against a model that assumes equal scale values, it shows that ambient color had a marginally significant effect on sweetness, $G^2(2) = 5.11$, $P = 0.078$. The u-parameters of the BTL model have ratio-scale properties; thus, it is meaningful to conclude that in the red illumination condition, the wine tasted about 50% sweeter than in either the blue or the white background color. As can be seen from the 95% confidence intervals in Fig. 11, the sweetness of the wine in red light was significantly higher than in blue or white light.

For the remaining attributes, the overall color effect was not significant, although pair-wise comparisons indicated a higher fruitiness in red than in blue or white light (see confidence intervals in Fig. 11).

For three attributes, there was a significant order effect favoring the first presentation interval. To illustrate, an estimated order effect of 0.50 means that the probability of choosing the first wine is twice as high as choosing the second, when they are presented under equally preferred illumination conditions. These findings are consistent with the rather rapid adaptation of taste and odor (e.g., von Békésy 1965; Zufall and Leinders-Zufall 2000).

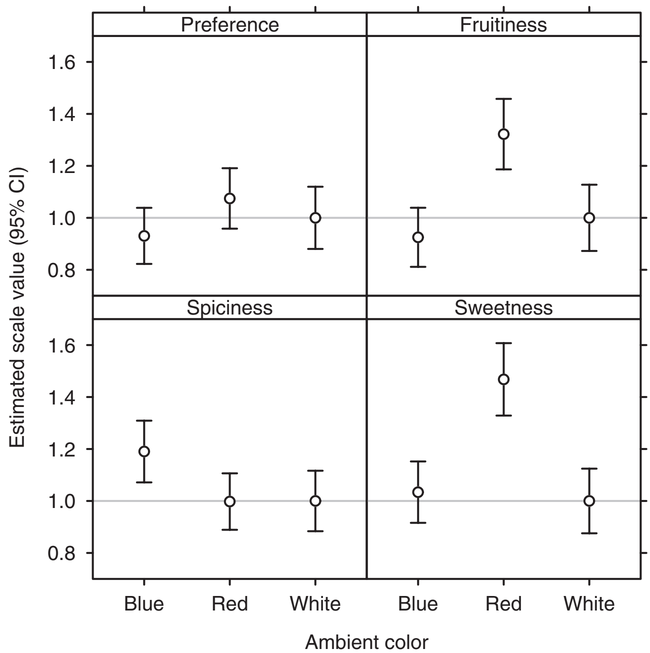

FIG. 11. EXPERIMENT 3: ATTRIBUTE SCALE VALUES ESTIMATED USING THE BRADLEY–TERRY–LUCE MODEL WITH ORDER EFFECT (Eq. 2)
For each attribute, white ambient color represents the unit of the scale. Error bars show 95% confidence intervals (CI).

GENERAL DISCUSSION

Ambient Color Exerts a Significant Albeit Small Effect on Flavor Judgments

When assessed between subjects at a local winery, blue and red room lighting prompted observers to like the Riesling wine that they tasted somewhat better compared with green or white lighting. The participants were also willing to spend more money on a bottle under blue and red lighting. Aroma differences could not be attributed to ambient color in this field experiment. However, when using more a sensitive within-subjects design in a laboratory environment, a number of interesting effects on aroma could be documented. The taste experience was clearly influenced by ambient color although merely sniffing the wine orthonasally was immune to lighting effects. Blue and green room lights made the wine appear spicier, although blue and red ambient color reduced the perceived fruitiness. Blue light also increased the perceived bitterness.

Our experiments demonstrate a significant effect of ambient lighting on flavor. This effect appears to be qualitatively different from the rather well-researched effects of beverage color on taste. The latter directly enter the holistic assessment of the object. The meaning of the object and its color form a unit to the extent that green bananas are hallucinated to be yellow (Hansen *et al.* 2006), and that false but typical aromas are attributed to objects whose colors have been manipulated (Morrot *et al.* 2001). These effects witness the importance of color adequacy or inadequacy (Zellner *et al.* 1991), but see (Zellner and Whitten 1999) and can be explained in a rather direct fashion of sensory-cognitive integration. Not so the effect of ambient color. The illumination of the room has no evolutionary and mostly no cultural correspondence with the beverages that are consumed in this room. This opens up the question as to why and how the effects of ambient color come about. One possible explanation for ambient color effects is the notion that ambient color changes the emotional or cognitive state of the observers, and thereby changes the setting for any flavor judgment. If this were the case, ambient color effects would be mediated and the research program should be to detect the paths of mediation.

Are Effects of Ambient Color Direct or Mediated Effects?

Experiment 2 failed to demonstrate a correlation between the emotional response to a color in terms of valence and arousal on the one hand and global liking of the wine tasted in this ambient color on the other hand. Thus, global liking of the wine is not a simple function of the color-induced emotion. For the flavor attributes spiciness, fruitiness and bitterness, for which ambient color had an effect in Experiment 2, there also appeared to be no simple relation between flavor and the valence or arousal associated with the color. For example, the wine tasted fruitier in green then in red or blue light, even though the valence of blue was higher than the valence of green, and the arousal of red was higher than the arousal of green. Taken together, the results speak against an emotion-mediated effect.

Generally, the rather rapid adaptation of taste poses a potentially serious problem to the assessment of color effects. We have all come across the experience that dry wine tastes rather horrible if sampled after a sweet late harvest. The significant order effects found in Experiment 3 demonstrate the importance of controlling for effects of stimulus presentation order in sensory studies. Although the stochastic model used for analyzing the paired-comparison data allowed separating order effects from effects of ambient color, it remains possible that the effects of ambient color was weakened by adaptation.

The three experiments taken together suggest that the effects of ambient light are rather genuine and render it unlikely that the effects be secondary to color inducing a particular emotional or adaptive state. This is not to say, however, that cognitive participation has played no role in the flavor judgments we solicited from our participants. For instance, we know that color affects cognitive processing, and just maybe, the effect that blue made the wine appear spicier was mediated by the positive effect that blue light has on alertness and cognitive processing (Revell *et al.* 2006).

ACKNOWLEDGMENTS

We are grateful to Judith Roßberg for organizing the experiment at the Allendorf winery, to Nevena Kalinova and Deuan Donev for their assistance in the data collection for Experiment 3, and to Agnes Münch for her assistance in technical issues.

REFERENCES

ALLEY, R.L. and ALLEY, T.R. 1998. The influence of physical state and color on perceived sweetness. J. Psychol. Interdiscip. Appl. *132*, 561–568.

AUBRY, V., ETIEVANT, P., SAUVAGEOT, F. and ISSANCHOU, S. 1999. Sensory analysis of Burgundy Pinot noir wines: A comparison of orthonasal and retronasal profiling. J. Sens. Studies *14*, 97–117.

BLAND, J.M. and ALTMAN, D.G. 1994. Correlation, regression, and repeated data. Br. Med. J. *308*, 896.

BLAND, J.M. and ALTMAN, D.G. 1995. Statistics Notes: 12. Calculating correlation-coefficients with repeated observations: Part 1 – correlation within subjects. Br. Med. J. *310*, 446–446.

BRADLEY, M.M. and LANG, P.J. 1994. Measuring emotion – the self-assessment mannequin and the semantic differential. J. Behav. Ther. Exp. Psychiatry *25*, 49–59.

BRADLEY, R.A. and TERRY, M.E. 1952. Rank analysis of incomplete block designs: I. The method of paired comparisons. Biometrika *39*, 324–345.

BURTON, P., GURRIN, L. and SLY, P. 1998. Extending the simple linear regression model to account for correlated responses: An introduction to generalized estimating equations and multi-level mixed modelling. Stat. Med. *17*, 1261–1291.

CAMGÖZ, N., YENER, C. and GUVENC, D. 2002. Effects of hue, saturation, and brightness on preference. Color Res. Appl. *27*, 199–207.

CHEN, D. and DALTON, P. 2005. The effect of emotion and personality on olfactory perception. Chem. Senses *30*, 345–351.

CLORE, G.L., SCHWARZ, N. and CONWAY, M. 1994. Affective causes and consequences of social information processing. In *Handbook of Social Cognition* (R.S. Wyer and T.K. Srull eds.) pp. 323–418, Lawrence Erlbaum, Hillsdale, NJ.

CLYDESDALE, F.M., GOVER, R. and FUGARDI, C. 1992. The effect of color on thirst quenching, sweetness, acceptability and flavor intensity in fruit punch flavored beverages. J. Food Qual. *15*, 19–38.

CROZIER, W.R. 1999. The meanings of colour: Preferences among hues. Pigment Resin Technol. *28*, 6–14.

DAVIDSON, R.R. and BEAVER, R.J. 1977. Extending the Bradley-Terry model to incorporate within-pair order effects. Biometrics *33*, 693–702.

DELWICHE, J. 2004. The impact of perceptual interactions on perceived flavor. Food Qual. Pref. *15*, 137–146.

DITTMAR, M. 2001. Changing colour preferences with ageing: A comparative study on younger and older native Germans aged 19–90 years. Gerontology *47*, 219–226.

DUBOSE, C.N., CARDELLO, A.V. and MALLER, O. 1980. Effects of colorants and flavorants on identification, perceived flavor intensity, and hedonic quality of fruit-flavored beverages and cake. J. Food Sci. *45*, 1393–1399.

FRANK, R.A., DUCHENY, K. and MIZE, S.J.S. 1989. Strawberry odor, but not red color, enhances the sweetness of sucrose solutions. Chem. Senses *14*, 371–377.

GUÉGUEN, N. 2003. The effect of glass colour on the evaluation of a Beverage's thirst-quenching quality. Curr. Psychol. Lett. *11*. Published online on 30 March 2006. http://cpl.revues.org/index398.html. Consulted on February 5, 2008.

HANSEN, T., OLKKONEN, M., WALTER, S. and GEGENFURTNER, K.R. 2006. Memory modulates color appearance. Nat. Neurosci. *9*, 1367–1368.

HERZ, R.S. and ENGEN, T. 1996. Odor memory: Review and analysis. Psychon. Bull. Rev. *3*, 300–313.

HOFFMANN, D., MÜLLER, S., SZOLNOKI, G., NÖLL, A. and SCHANOWSKI, B. 2006. Präferenzen für Weissweine aus Rheinland-Pfalz [Consumer preferences for white wine from Rhineland-Palatine]. In *Weißweinstudie. Repräsentative Verbraucheruntersuchung: Zielgruppen und Präferenzen* (Ministerium für Wirtschaft, Verkehr, Landwirtschaft und Weinbau Rheinland-Pfalz, ed.) pp. 5–91, Ministerium für Wirtschaft, Verkehr, Landwirtschaft und Weinbau Rheinland-Pfalz, Mainz.

HYMAN, A. 1983. The influence of color on the taste perception of carbonated water preparations. Bull. Psychon. Soc. *21*, 145–148.

JACKSON, R.S. 2002. *Wine Tasting: A Professional Handbook.* Academic Press, San Diego, CA.

KAISER, P.K. 1984. Physiological response to color – A critical-review. Color Res. Appl. *9*, 29–36.

KEMP, S.E. and GILBERT, A.N. 1997. Odor intensity and color lightness are correlated sensory dimensions. Am. J. Psychol. *110*, 35–46.

KESELMAN, H.J., ALGINA, J. and KOWALCHUK, R.K. 2001. The analysis of repeated measures designs: A review. Br. J. Math. Stat. Psychol. *54*, 1–20.

KOCH, C. and KOCH, E.C. 2003. Preconceptions of taste based on color. J. Psychol. Interdiscip. Appl. *137*, 233–242.

KOZA, B.J., CILMI, A., DOLESE, M. and ZELLNER, D.A. 2005. Color enhances orthonasal olfactory intensity and reduces retronasal olfactory intensity. Chem. Senses *30*, 643–649.

LANG, P.J. 1980. Behavioral treatment and bio-behavioral assessment: Computer applications. In *Technology in Mental Health Care Delivery Systems*, (J.B. Sidowski, J.H. Johnson and T.A. Williams, eds.) pp. 119–137, Ablex, Norwood, NJ.

LIANG, K.Y. and ZEGER, S.L. 1986. Longitudinal data analysis using generalized linear models. Biometrika *73*, 13–22.

LIANG, K.Y. and ZEGER, S.L. 1993. Regression analysis for correlated data. Annu. Rev. Public Health *14*, 43–68.

LOVIE, P. 1986. Identifying outliers. In *New Developments in Statistics for Psychology and the Social Sciences*, (A.D. Lovie, ed.) pp. 45–69, The British Psychological Society, Leicester, UK.

LUCE, R.D. 1959. *Individual Choice Behavior; A Theoretical Analysis.* Wiley, New York.

LUKAS, J. 1991. BTL-Skalierung verschiedener Geschmacksqualitäten von Sekt (BTL-scaling of different taste qualities of champagne). Z. Exp. Angew. Psychol. *38*, 605–619.

MCCULLAGH, P. 1980. Regression-models for ordinal data. J. R. Stat. Soc. Ser. B Methodol. *42*, 109–142.

MEHRABIAN, A. and RUSSELL, J.A. 1974. *An Approach to Environmental Psychology*. M.I.T. Press, Cambridge.

MORROT, G., BROCHET, F. and DUBOURDIEU, D. 2001. The color of odors. Brain Lang. *79*, 309–320.

MOSER, B.K. and STEVENS, G.R. 1992. Homogeneity of variance in the two-sample means test. Am. Stat. *46*, 19–21.

OSGOOD, C.E., SUCI, G.J. and TANNENBAUM, P.H. 1957. *The Measurement of Meaning*. Univer. Illinois Press, Oxford, England.

ÖSTERBAUER, R.A., MATTHEWS, P.M., JENKINSON, M., BECKMANN, C.F., HANSEN, P.C. and CALVERT, G.A. 2005. Color of scents: Chromatic stimuli modulate odor responses in the human brain. J. Neurophysiol. *93*, 3434–3441.

OU, L.C., LUO, M.R., WOODCOCK, A. and WRIGHT, A. 2004. A study of colour emotion and colour preference. Part III: Colour preference Modeling. Color Res. Appl. *29*, 381–389.

OVERALL, J.E. and TONIDANDEL, S. 2004. Robustness of generalized estimating equation (GEE) tests of significance against misspecification of the error structure model. Biom. J. *46*, 203–213.

PARR, W.V., HEATHERBELL, D. and WHITE, K.G. 2002. Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. Chem. Senses *27*, 747–755.

PENDERGAST, J.F., GANGE, S.J., NEWTON, M.A., LINDSTROM, M.J., PALTA, M. and FISHER, M.R. 1996. A survey of methods for analyzing clustered binary response data. Int. Stat. Rev. *64*, 89–118.

REVELL, V.L., ARENDT, J., LOUIS, F.F. and SKENE, D.J. 2006. Alerting effects of light are sensitive to very short wavelengths. Neurosci. Lett. *399*, 96–100.

RUSSELL, J.A. 1980. A circumplex model of affect. J. Pers. Soc. Psychol. *39*, 1161–1178.

SAUVAGEOT, F. and STRUILLOU, A. 1997. Effet d'une modification de la couleur des échantillons et de l'éclairage sur la flaveur de vins évaluée sur une échelle de similarité (Effect of the modification of wine colour and lighting conditions on the perceived flavour of wine, as measured by a similarity scale). Science des Aliments *17*, 45–67.

SCHLOSBERG, H. 1952. The description of facial expressions in terms of two dimensions. J. Exp. Psychol. *44*, 229–237.

SCHLOSBERG, H. 1954. Three dimensions of emotion. Psychol. Rev. *61*, 81–88.

SMALL, D.M., VOSS, J., MAK, Y.E., SIMMONS, K.B., PARRISH, T. and GITELMAN, D. 2004. Experience-dependent neural integration of taste and smell in the human brain. J. Neurophysiol. *92*, 1892–1903.

SUK, H.-J. 2006. *Color and emotion. A study on the affective judgment of color across media and in relation to visual stimuli*. Unpublished dissertation. Universität Mannheim, Mannheim.

TERWOGT, M.M. and HOEKSMA, J.B. 1995. Colors and emotions – preferences and combinations. J. Gen. Psychol. *122*, 5–17.

VALDEZ, P. and MEHRABIAN, A. 1994. Effects of color on emotions. J. Exp. Psychol. Gen. *123*, 394–409.

VON BÉKÉSY, G. 1965. The effect of adaptation on the taste threshold observed with a semiautomatic gustometer. J. Gen. Physiol. *48*, 481–488.

WHITFIELD, T.W. and WILTSHIRE, T.J. 1990. Color psychology: A critical review. Genet. Soc. Gen. Psychol. Monogr. *116*, 385–411.

WICKELMAIER, F. and SCHMID, C. 2004. A Matlab function to estimate choice model parameters from paired-comparison data. Behav. Res. Methods Instrum. Comput. *36*, 29–40.

WUNDT, W.M. 1896. *Grundriss der Psychologie.* W. Engelmann, Leipzig.

WYSZECKI, G. and STILES, W.S. 1982. *Color Science: Concepts and Methods, Quantitative Data and Formulae.* Wiley, New York.

ZAMPINI, M., SANABRIA, D., PHILLIPS, N. and SPENCE, C. 2007. The multisensory perception of flavor: Assessing the influence of color cues on flavor discrimination responses. Food Qual. Pref. *18*, 975–984.

ZELLNER, D.A. and DURLACH, P. 2003. Effect of color on expected and experienced refreshment, intensity, and liking of beverages. Am. J. Psychol. *116*, 633–647.

ZELLNER, D.A. and KAUTZ, M.A. 1990. Color affects perceived odor intensity. J. Exp. Psychol. Hum. Percept. Perform. *16*, 391–397.

ZELLNER, D.A. and WHITTEN, L.A. 1999. The effect of color intensity and appropriateness on color-induced odor enhancement. Am. J. Psychol. *112*, 585–604.

ZELLNER, D.A., BARTOLI, A.M. and ECKARD, R. 1991. Influence of color on odor identification and liking ratings. Am. J. Psychol. *104*, 547–561.

ZORN, C.J.W. 2001. Generalized estimating equation models for correlated data: A review with applications. Am. J. Polit. Sci. *45*, 470–490.

ZUFALL, F. and LEINDERS-ZUFALL, T. 2000. The cellular and molecular basis of odor adaptation. Chem. Senses *25*, 473–481.