Chem. Senses 38: 305–313, 2013

doi:10.1093/chemse/bjs142
Advance Access publication January 16, 2013

Identification of Gustatory–Olfactory Flavor Mixtures: Effects of Linguistic Labeling

Jennifer M. Brewer[1], Adam Y. Shavit[1], Timothy G. Shepard[1], Maria G. Veldhuizen[1,2], Roshan Parikh[1] and Lawrence E. Marks[1,3,4]

[1]The John B. Pierce Laboratory, 290 Congress Avenue, New Haven, CT 06519, USA, [2]Department of Psychiatry, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA, [3]Department of Environmental Health Sciences, Yale School of Public Health, 60 College Street, New Haven, CT 06520, USA and [4]Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06510, USA

Correspondence to be sent to: Lawrence E. Marks, The John B. Pierce Laboratory, 290 Congress Avenue, New Haven, CT 06519, USA. e-mail: lmarks@jbpierce.org

Accepted December 18, 2012

Abstract

Two experiments, using different ranges and numbers of stimuli, examined how linguistic labels affect the identification of flavor mixtures containing different proportions of sucrose (gustatory flavorant) and citral (olfactory flavorant). Both experiments asked subjects to identify each stimulus as having either "mostly sugar" or "mostly citrus." In one condition, no labels preceded the flavor stimuli. In another condition, each flavor stimulus followed a label, either SUGAR or CITRUS, which, the subjects were informed, usually though not always named the stronger flavor component; that is, the labels were probabilistically valid. The results of both experiments showed that the labels systematically modified the identification responses: Subjects responded "sugar" or "citrus" more often when the flavor stimulus followed the corresponding label, SUGAR or CITRUS. But the labels hardly affected overall accuracy of identification. Accuracy was possibly limited, however, by both the confusability of the flavor stimuli per se and the way that confusability could limit the opportunity to discern the probabilistic associations between labels and individual flavor stimuli. We describe the results in terms of a decision-theoretic model, in which labels induce shifts in response criteria governing the identification responses, or possibly effect changes in the sensory representations of the flavorants themselves.

Key words: flavor, gustatory, identification, labeling, mixtures, olfactory

Introduction

Flavors signal the identity and composition of foods and drinks and indicate their potential harm or benefit. Flavor thereby plays an important role in the consumption of food and drinks, thus in energy balance, fluid balance, and, ultimately, body weight. Although sometimes perceived as unitary experiences, flavors nevertheless reflect the output of a dynamic system that processes signals from gustation, olfaction, and somatosensation (Small and Prescott 2005), even vision and hearing (Auvray and Spence 2008), then combines the multisensory information with knowledge and expectations developed through both recent and long-term perceptual experiences (Small et al. 2004; Koza et al. 2005).

In the present study, we ask how one kind of cognitive information, the linguistic label given just before the subject samples each flavorant, affects the identification of that flavorant. Both past experience and current context typically lead to expectations about the foods and beverages we choose to eat and drink. When we are about to consume a soft drink, for example, we generally know in advance whether it will be, say, a cola; and if we also know the brand of cola, then we may well expect the flavor to have a particular level of sweetness or the presence of perhaps a vanilla note. Prior knowledge can affect both expectations before eating or drinking and judgments made after tasting, as evident in the following examples.

Kähkönen and Tuorila (1998) asked how linguistic information (verbal labels) can affect sensory and hedonic ratings of full-fat and reduced-fat Bologna sausage. When given only the labels but not the foods themselves, subjects expected that reduced-fat sausage would be less salty and less fatty than regular sausage; after tasting the 2 sausages, the ratings still differed, although the differences were smaller. When another group of subjects sampled the 2 sausages without the labels, those subjects judged the sausages to be similar. To the subjects given the labels, therefore, the judgments of sensory properties of the sausage presumably represent a compromise between the sensory information and the linguistically produced information. In a similar vein, Tuorila et al. (1994) showed how incorrect and correct labels of fat-free and regular-fat pound cake, crackers, and American cheese could affect sensory and hedonic expectations.

Using a novel food stimulus, salmon-flavored ice cream, which could be labeled as either ice cream or savory mousse, Yeomans et al. (2008) found substantial interactions between sensory information and linguistic information. Labeling the food as ice cream rather than frozen savory mousse decreased the ratings of pleasantness but increased the ratings of saltiness and overall flavor intensity. Given the findings from a separate experiment that labeling the food as ice cream generated expectations of a sweet flavor, the ratings likely reflected contrast between the label-induced expectations and the sensory flavor signals. In yet another experiment, subjects who saw the food but did not taste it rated the food as sweeter when it was labeled as ice cream and saltier when it was labeled as mousse. Again, the judgments likely reflected an interaction between expectations, generated by the labels, and the sensory flavor information.

These findings indicate that linguistic information about a food sets up expectations that can then interact with sensory information obtained from sampling the food itself. Although interactions may sometimes be contrastive (e.g., Yeomans et al. 2008), judgments commonly suggest assimilation—the integration (addition) of linguistic information and sensory information (e.g., Tuorila et al. 1994; Kähkönen and Tuorila 1998; Wansink et al. 2005).

The present experiment examined how linguistic information interacts with sensory information, capitalizing on the use of a simple model system of flavorants: artificial beverages containing different proportions of 2 flavorants, the sweet-tasting gustatory flavorant sucrose and the citrus-"tasting" olfactory flavorant citral. One helpful feature of using these gustatory–olfactory mixtures is the evidence that the perceived intensities of sucrose and citral combine more or less additively (Murphy and Cain 1980; Marks et al. 2012). Although not crucial to the experiment, the approximate additivity suggests that there is relatively little interaction between sucrose and citral, at least with regard to perceived intensity. More importantly, by varying the proportions of sucrose and citral, we can produce a set of flavor stimuli that vary perceptually from predominantly sweet (having relatively more sucrose) to predominantly citrus (having relatively more citral). Such a stimulus set lends itself well to a study of identification in which the subject's task on each trial is to identify the dominant flavor component. We can then compare the identification functions obtained when the flavor stimuli follow different linguistic labels. Lawless et al. (1991) created an analogous set of odor stimuli by varying the proportions of 2 odorants having different perceived qualities. Here, we adapt the stimuli of Lawless et al. to gustatory–olfactory flavor mixtures.

The studies reviewed above imply that information about a food stimulus that is provided prior to sampling the stimulus can modify the overt response, possibly because the information elicits expectations that, in turn, may modify the subsequent perception or at least modify the overt response. The tendency to change expectations on the basis of prior information comes, no doubt, from experiences in which the information is valid, at least probabilistically. Indeed, this is likely the basis for learning the referents for words such as "sweet" and "salty." If the label SUGAR precedes flavorants that usually contain more sucrose than citral, and CITRUS precedes flavorants that usually contain more citral, then subjects may pick up this information and use it to improve the accuracy of identification. It is also possible, however, that labels simply increase the probability that responses will match the labels. The present pair of experiments may shed light on which of these processes (or both) underlies the responses of subjects to a well-defined set of gustatory–olfactory flavorants. Both experiments used the same method, but examined effects of labeling on different sets of stimulus mixtures.

Materials and methods

Subjects

A total of 46 men and women participated, 16 in the initial Experiment 1 and 30 in a more extensive Experiment 2 (no subject served in both). Following a preliminary session, which assessed the perceived flavor intensity of sucrose and citral (all subjects in both experiments), each subject then participated in 1 identification session (Experiment 1) or 2 identification sessions (Experiment 2), held on separate days. Experiment 1 tested 8 subjects in a baseline (no label) condition [6 women and 2 men, 22–32 years: mean = 27.3; standard deviation (SD) = 4.0] and the other 8 in a labeling condition (7 women and 1 man, 22–31 years: mean = 25.3; SD = 3.1). Experiment 2 tested 12 subjects in a baseline condition (8 women and 4 men, 19–45 years: mean = 29.0, SD = 7.2) and the other 18 in a labeling condition (12 women and 6 men, 18–29 years: mean = 22.9, SD = 3.3). The subjects did not report any taste impairments, and they were instructed not to eat or drink anything but water in the hour before the sessions. Subjects were paid $10/hour to participate. Most of the subjects were affiliated with Yale University. The research complied with the *Declaration of Helsinki* for Medical Research involving Human Subjects. All subjects gave informed consent, under a protocol approved by Yale University's Human Subjects Committee.

Materials

Stimuli were made fresh at least every week and refrigerated until warmed to room temperature (23.5 °C) before testing. The gustatory flavorant was sucrose (J.T. Baker, CAS# 57-50-1, $C_{12}H_{22}O_{11}$) dissolved in deionized water. The olfactory stimulus was citral (International Flavors and Fragrances, CAS# 5392-40-5, chemical characterization: 3,7-dimethyl-2,6-octadienal, a mixture of *cis*- and *trans*-isomers). Because citral does not dissolve readily in water, we first dissolved the citral in 200 proof ethyl alcohol (ethanol, CAS# 64-17-5), in a ratio of 1 part citral to 30 parts ethanol, and then added deionized water to bring the concentrations of citral to the desired values. Concentrations of ethanol and citral were below trigeminal and taste thresholds (Wilson et al. 1973; Cometto-Muniz and Cain 1990; Cerf-Ducastel and Murphy 2001).

The preliminary sessions of each experiment presented 6 concentrations each of sucrose (0.008–0.099 M) and of citral (6.6×10^{-5}–1.09×10^{-1} M). Based on the results of these sessions, we chose the concentration of citral that gave a comparable rating of perceived intensity to 0.099 M sucrose across subjects, which was 0.0059 M citral. Then, for each condition (no labels or labeling) of each experiment, we constructed either 7 mixtures (Experiment 1) or 4 mixtures (Experiment 2) containing the following relative proportions of the flavorant concentrations: Experiment 1: 0.10 sucrose/0.90 citral, 0.25 sucrose/0.75 citral, 0.40 sucrose/0.60 citral, 0.50 sucrose/0.50 citral, 0.60 sucrose/0.40 citral, 0.75 sucrose/0.25 citral, and 0.90 sucrose/0.10 citral; Experiment 2: 0.35 sucrose/0.65 citral, 0.45 sucrose/0.55 citral, 0.55 sucrose/0.45 citral, and 0.65 sucrose/0.35 citral. Thus, for example, the mixture denoted as 0.35 sucrose/0.65 citral contained 0.035 M sucrose (0.099×0.35) and 0.0038 M citral (0.0059×0.65).

Procedure

Preliminary session

The procedure used in the preliminary session was the same for all subjects. Subjects judged the 6 concentrations of sucrose and the 6 concentrations of citral in separate blocks of trials, each stimulus being presented 10 times within the block in random order (60 presentations each of sucrose and citral). Subjects rated the perceived intensity of each flavor stimulus on a Labeled Magnitude Scale (Green et al. 1996), as modified by Marks et al. (2012). This modified scale contains numbers as well as verbal descriptors, allowing the subjects to respond with a number to represent the intensity of the flavorant based on the number's location relative to the descriptors. In the preliminary and main sessions, each trial presented 5 mL of flavor solution in a 30-mL plastic cup. Subjects rinsed thoroughly with deionized water before each stimulus. The next stimulus was presented about 30 s after the judgment of the previous stimulus. Each block of 60 trials in the preliminary session took about 15 min. The ratings obtained at baseline showed that 0.099 M sucrose was judged about as strong as 0.0059 M citral, and these concentrations served as the stimuli in the main experiments.

Identification experiments

In the identification experiments, each of the 7 flavor mixtures (Experiment 1) or 4 flavor mixtures (Experiment 2) was presented a total of 12 times or 20 times, respectively, within a session in random order. Thus, each subject in Experiment 1 provided 84 identifications in all, and each subject in Experiment 2 provided 160 (80 in each of the 2 sessions). The task on each trial was to identify the dominant component of the stimulus, "sugar" or "citrus." The no-label condition served as a baseline, measuring how, in the absence of labels, the proportions of "sugar" and "citrus" responses vary with the proportions of the 2 flavorants in the absence of labels.

In the label condition, each flavor stimulus followed a linguistic label, either SUGAR or CITRUS, presented on a computer monitor. The subjects were informed that on most trials, but not all of them, the label would correctly denote the dominant flavor. Although the label would not indicate the dominant flavor on some of the trials, the labels would be informative overall. Subjects were not informed, however, as to the specific statistical associations between the labels and the flavor stimuli. In Experiment 1, the labels SUGAR/CITRUS were assigned to the 7 flavor stimuli, in increasing proportion of sucrose, with the frequencies 0/12, 0/12, 3/9, 6/6, 9/3, 12/0, and 12/0. In Experiment 2, the labels SUGAR/CITRUS were assigned to the 4 flavor stimuli, in increasing proportion of sucrose, with the frequencies 4/16, 8/12, 12/8, and 16/4. Thus, in Experiment 1, the 4 most extreme stimuli were always labeled correctly, either as CITRAL (0.10S/0.90C and 0.25S/0.75C) or as SUGAR (0.75S/0.25C and 0.90S/0.10C), and only the 3 intermediate stimuli (0.40S/0.60C, 0.50S/0.50C, and 0.60S/0.40C) were labeled probabilistically. In Experiment 2, however, all 4 stimuli were labeled probabilistically. We chose these frequencies of presentation to ensure that the labels would be statistically associated with the dominant components of the stimulus ensembles in both experiments while avoiding obvious mismatches between labels and flavors when the proportion of sucrose to citral was either very small or very large in Experiment 1. Overall, the validity of the labels (the proportion of trials in which the labels denoted the correct responses) was greater in Experiment 1 than in Experiment 2 [0.86 vs. 0.70; in Experiment 1, we defined the validity of the labels given to the middle stimulus of the series (0.50S/0.50C) as 0.5].

Results and discussion

Effects of labeling on "sugar" and "citrus" responses

Figure 1 shows the results obtained in the baseline (no-label) conditions of both experiments, plotting the average proportion of "sugar" responses as a function of the proportion of sucrose to citral in the stimulus. In both experiments, the proportion of "sugar" responses was at or near 0.5 when the stimulus contained equal relative proportions of sucrose and citral, consistent with the expectation that subjects would be equally likely to identify a flavorant as "sugar" and "citrus" when the components were equated in perceived intensity.

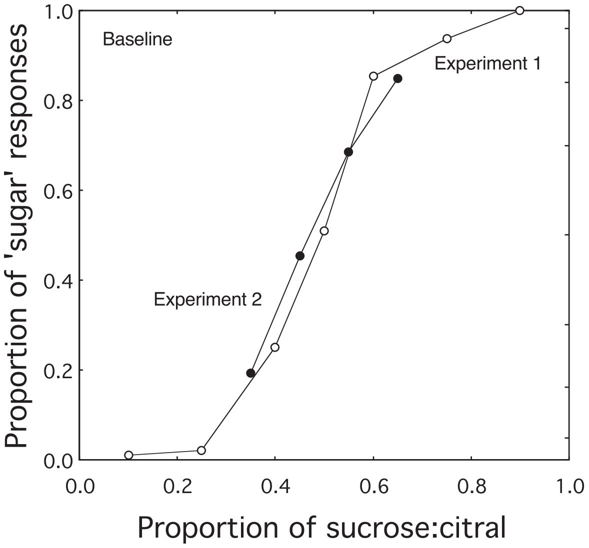

Figure 1 Proportion of "sugar" responses, plotted as a function of the relative proportion of sucrose to citral in Experiment 1 (open circles) and Experiment 2 (filled circles), when no labels preceded the flavor stimuli (baseline condition).

Figure 2 shows the analogous results obtained in the labeling conditions, with the results from Experiments 1 and 2 given in the figure's left and right panels, respectively. Each panel again plots the average proportion of "sugar" responses as a function of the proportion of sucrose in the flavor stimulus, separately now for those trials in which the flavorant was labeled SUGAR and for those trials in which it was labeled CITRUS. The main statistical analyses (analyses of variance with repeated measures) were conducted on responses to all of the stimuli that received both labels: the middle 3 stimuli of the series in Experiment 1 and all 4 stimuli in Experiment 2.

Not surprisingly, the proportion of "sugar" responses increased significantly with increasing proportion of sucrose in the stimulus: Experiment 1, $F(2,14) = 36.8$, $P < 0.0001$; Experiment 2, $F(3,51) = 155.1$, $P < 0.0001$. Most importantly, in both experiments, the label markedly affected identification: The proportion of "sugar" responses was greater when the label was SUGAR rather than CITRUS, the effect of labeling being significant in both Experiment 1, $F(1,7) = 11.8$, $P = 0.01$ and Experiment 2, $F(1,17) = 11.6$, $P < 0.0035$. Further, in both experiments, the 2 identification functions are displaced more or less uniformly, implying that shifting the labels exerted a roughly constant effect on responses to all of the stimuli—at least over the limited stimulus ranges used with both labels. The interaction between label and stimulus was not significant in either Experiment 1, $F(2,14) < 1$ or Experiment 2, $F(3,51) = 1.10$, $P > 0.35$. Moreover, in both experiments, the average proportion of "sugar" responses was greater when the flavor stimuli were labeled SUGAR (0.70 and 0.58 in Experiments 1 and 2, respectively), smaller when the stimuli were labeled CITRUS (0.45 and 0.44). By way of comparison, proportions were intermediate at baseline, when labels were absent (0.54 and 0.55).

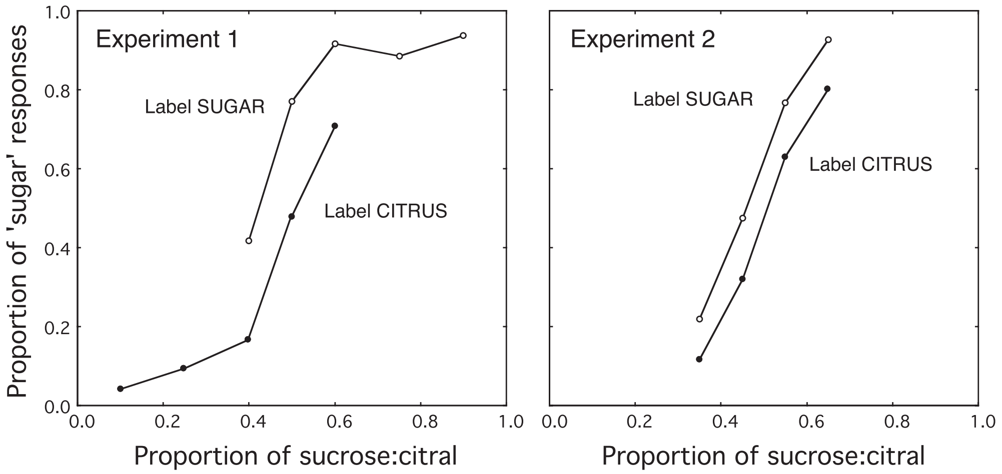

Figure 2 Proportion of "sugar responses" in the labeling conditions of Experiment 1 (left panel) and Experiment 2 (right panel), plotted separately for trials in which the flavor stimulus was preceded by the labels SUGAR (open circles) and CITRUS (filled circles).

Finally, note that labeling a flavorant as SUGAR rather than CITRUS led to an average increase of 0.25 in the probability of a "sugar" response in Experiment 1, but to an average increase of only 0.14 in Experiment 2. Although the effect of labeling was numerically greater in Experiment 1, the difference is not significant, $t(24) = 1.37$, $P = 0.18$.

Effects of labeling on accuracy

We defined responses of "sugar" as correct when the proportion of sucrose in the stimulus exceeded 0.5, and defined responses of "citrus" as correct when the proportion of citral exceeded 0.5. Then, we calculated the proportions of correct responses both with and without the labels (in Experiment 1, we defined as correct half of all responses to the ambiguous, 0.50S/0.50C stimulus). In the absence of labels, the proportions of correct responses presumably characterize the contribution to identification of the sensory flavor system by itself; with the addition of the labels, the corresponding proportions presumably characterize the joint contributions of the chemosensory system and linguistically based processes. The left and right panels of Figure 3 show how accuracy varied with stimulus proportion in the no-label (baseline) and labeling conditions of Experiments 1 and 2, respectively.

In Experiment 1, responses to 2 of the stimuli, 0.40S/0.60C and 0.60S/0.40C, are most important, and this is so for 2 reasons: First, each stimulus could be unambiguously defined as primarily citrus or sugar. And second, each was presented in all 3 conditions—no label, label CITRUS, and label SUGAR. The results obtained with these 2 stimuli are simple and clear: Accuracy was best when the label was valid (CITRUS with the flavor stimulus having more citral, SUGAR with the stimulus having more sucrose) and worst when the label was invalid (SUGAR with more citral, CITRUS with more sucrose). Accuracy in the absence of labels fell between. This outcome follows directly from the general tendency, evident in Figure 2, for SUGAR labels to lead to higher proportions of "sugar" responses and CITRUS labels to lead to higher proportions of "citrus" responses.

Labeling in Experiment 1 had a different effect, however, on accuracy in identifying the extreme stimuli. When the flavor stimulus contained mostly citral (0.10S/0.90C and 0.25S/0.75C), accuracy was greater in the absence of labels than it was following the label CITRUS (CITRUS always being the label associated with these stimuli). The outcome was analogous at the other end of the stimulus scale (0.75S/0.25C and 0.90S/0.10C), where the only label associated with these stimuli, SUGAR, was always valid, but accuracy was nevertheless again better when the label was absent. This outcome is intriguing: It suggests that the sensory signals, which in the absence of labels led to virtually perfect accuracy, failed to do so in the presence of labels, even though the labels should reinforce correct responding—perhaps because the presence of labels in the session induced uncertainty. By implication, these results imply that the subjects did not distinctly associate the stimulus-specific probabilities of the labels with each flavor stimulus.

In Experiment 2, the labels appeared to affect accuracy in different ways when the flavor stimuli contained relatively more sucrose (0.55S/0.45C and 0.65S/0.35C) and when the stimuli contained relatively more citral (0.35S/0.65C and 0.45S/0.55C). When the flavor stimuli had mostly sucrose, the label SUGAR increased accuracy relative to baseline, whereas the label CITRUS decreased accuracy, much as in Experiment 1. When the flavor stimuli had mostly citral, however, the label CITRUS increased accuracy relative to baseline, but SUGAR did not decrease accuracy at all. This last result could of course be spurious, and it is worth noting that, in Experiment 1, labeling the mostly citral stimulus as SUGAR did clearly decrease accuracy.

Save for the anomalous results in Experiment 2 with stimuli containing mostly citral, the measures of accuracy in the 2 experiments are broadly compatible with the general observation that labeling the stimuli as CITRUS or

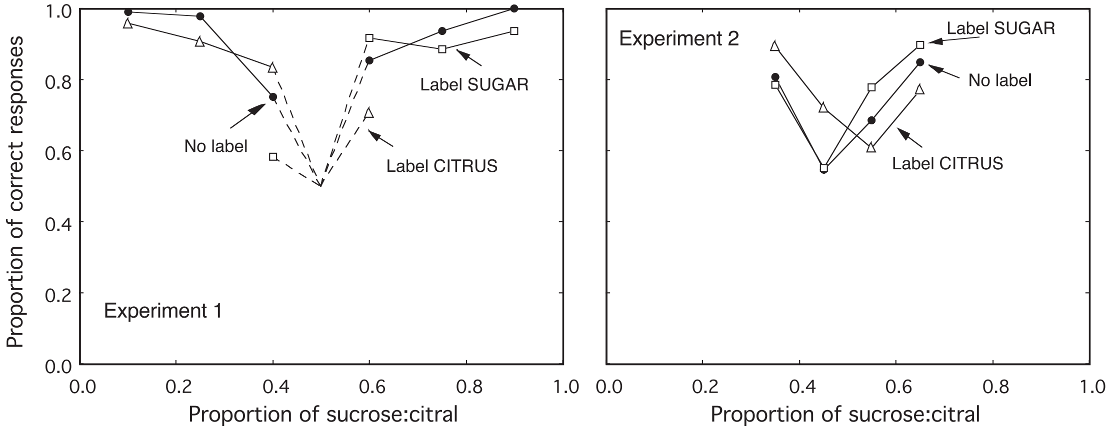

Figure 3 Proportion of correct responses in the no-label condition (filled circles), label-CITRUS condition (open triangles), and label-SUGAR condition (open squares) of Experiment 1 (left panel) and Experiment 2 (right panel), plotted as a function of the proportion of sucrose to citral. In Experiment 1, 50% of the responses to the stimulus containing a sucrose:citral proportion of 0.5 were treated as correct.

SUGAR increased the tendency to identify the stimulus as "citrus" or "sugar," respectively. Because this tendency increases the accuracy of identification when the labels are valid but decreases accuracy when the labels are invalid, the overall effect of labeling on accuracy should be modest. This is indeed the case when we calculate overall measures of accuracy with and without labels. In Experiment 1, overall accuracy was actually worse when the flavor stimuli were presented with labels than without them, 0.83 versus 0.86, although the difference is small and not statistically reliable, $t(14) = 0.70$, $P > 0.45$ (omitting responses to the 0.50S/0.50C stimulus increases the averages to 0.89 and 0.92, the difference still being nonsignificant, $t(14) = 0.74$, $P > 0.45$). In Experiment 2, however, accuracy was significantly greater with labels than without them, 0.78 versus 0.72, $t(28) = 2.49$, $P < 0.02$.

Finally, we note that accuracy in Experiment 1 surpassed that in Experiment 2. This is not surprising, given that Experiment 1 used a wider range of stimulus proportions, with larger differences between successive stimuli. Thus, the overall rate of correct responding was significantly greater in Experiment 1 than in Experiment 2 in the absence of labels (0.86 vs. 0.72, $t(19) = 6.01$, $P < 0.0001$) and was greater though not significantly so when the flavor stimuli followed labels (0.83 vs. 0.78, $t(24) = 1.88$, $P = 0.072$).

General discussion

Two experiments examined how linguistic labels affected the identification of flavor stimuli containing different proportions of the gustatory flavorant sucrose and the olfactory flavorant citral, asking subjects to identify each stimulus as containing either "mostly sugar" or "mostly citrus." In both experiments, labeling a flavor stimulus as SUGAR rather than CITRUS led to more responses of "sugar" than of "citrus." Consequently, the label SUCROSE tended to increase the accuracy of identification when the flavor stimulus contained more sucrose than citral, but to decrease accuracy when the stimulus contained more citral than sucrose. Similarly, the label CITRUS tended to increase or decrease accuracy when flavor stimulus contained proportionally more or less citral than sucrose. Thus, some of the label-induced responses were correct, some incorrect.

Although the labels themselves were statistically informative overall, subjects nevertheless relied substantially on the chemosensory signals. As is clear in Figure 2, the effect on identification of varying the proportion of sucrose in the flavor stimulus substantially exceeded the effect of varying the linguistic label. Further, in both Experiments 1 and 2, the functions obtained with the labels SUGAR and CITRUS are roughly parallel, indicating that the difference between the effects of the 2 labels was relatively constant across the various flavor stimuli within each experiment—although the difference probably would likely have diminished had we presented both labels with flavor stimuli at the extremes of the psychometric function, where performance in the absence of labels becomes asymptotic.

In brief, the present results imply that flavor identification can reflect an integration or interaction of sensory information from flavor stimuli with linguistic information that presumably creates expectations regarding the identity of the stimulus. The results complement findings reported by other investigators on the ways that verbal information affected judgments of fat-free versus regular-fat foods (Tuorila et al. 1994), yogurts (Schifferstein et al. 1999), entrees on menus at a cafeteria (Wansink et al. 2005), frozen mousse (Yeomans et al. 2008), and odors of various foods (Distel and Hudson 2001). Our results extend these findings by showing how a label can systematically influence the identification of the primary component of gustatory–olfactory flavor mixtures. In the next 2 sections, we consider further just how the sensory and linguistic information may combine.

Validity of labels and accuracy of identification

The labels given in both of the present experiments were probabilistically valid; that is, each label usually but not always named the dominant component of the subsequent flavor stimulus. Had the labels always correctly named the subsequent flavors, the subjects might have ignored the flavor percepts and simply responded by echoing the labels. On the other hand, had the labels been randomly associated with the flavor stimuli, the subjects might have ignored the labels and relied exclusively on the sensory flavor signals. Given our experimental design, the labels were valid (named the dominant component) on 86% of the trials in Experiment 1 and on 70% of the trials in Experiment 2. If the subjects had ignored the flavor information and simply responded by repeating the labels, performance in both experiments would have exceeded chance—and the subjects could have done this even if they did not discern the specific statistical associations of the 2 labels with each flavor stimulus.

In Experiment 1, the labels were always valid when they preceded the 4 stimuli that were most easily identified: CITRUS always preceded stimuli having citral proportions of 0.75 and 0.90, and SUGAR always preceded stimuli having sucrose proportions of 0.75 and 0.90. These 4 stimuli were almost always identified correctly in the baseline, no-label condition, but were identified less than perfectly when they followed labels, even though the labels were always valid. This outcome, perhaps counterintuitive, suggests that the subjects were sensitive to global associations between the labels and the stimuli—CITRUS being associated mostly with flavors that seemed more "citrusy," SUGAR associated mostly with flavors that seemed more "sugary"—and not very sensitive to the unique associations between the labels and each specific stimulus. Perhaps, the flavor stimuli were not sufficiently discriminable from one another to allow the subjects to determine how the 2 labels associated with each flavor stimulus. In Experiment 1, the overall validity of the labels, 0.86, slightly exceeded the subjects' accuracy with labels, 0.83; in this experiment, subjects would actually have been slightly more accurate overall had they simple repeated the label on each trial. In Experiment 2, however, accuracy with labels (0.78) did exceed the validity of the labels (0.70), as well as accuracy in the absence of labels (0.72). In Experiment 2, therefore, the subjects presumably were able to discern to some extent the specific associations between labels and flavor stimuli, using the labels to enhance the information available from the chemosensory signals.

That the subjects continued to rely substantially on the chemosensory signals even when given labels is evident in Figure 3, which shows the accuracy of responding to each stimulus in each condition of labeling. Note that if subjects had relied exclusively on the labels, then, on all of the trials on which the labels were not valid—SUCROSE preceding stimuli having proportions of sucrose >0.5, CITRUS preceding stimuli having proportions of sucrose >0.5—accuracy would have been zero. In fact, accuracy was always considerably greater than zero, indeed, always greater than chance (0.5) for every flavor stimulus with each label, as it was when the flavor stimuli were presented without labels and the subjects had to rely exclusively on the chemosensory signals. In Experiment 2 at least, subjects presumably used information in the labels, without substantially misusing it.

Effects of labeling: integration and interference?

From the pattern of results, especially Figure 2, we discern 2 distinct effects of the labels on flavor identification. The first effect is assimilation, and this is evident when we examine how the labels, SUGAR and CITRUS, differentially affected responding: With labeling, the responses tend to resemble the labels, with SUGAR leading to a greater propensity to respond "sugar" and CITRUS leading to a greater propensity to respond "citrus." The second effect is interference, and this is evident when we compare performance with and without labels, especially in Experiment 1: The results of Experiment 1 imply that labels can sow confusion. Adding labels, even when always valid, nevertheless impaired performance in identifying those 4 stimuli that were most readily identified without the labels. From this outcome, we infer that the interference resulted not from the (valid) labels per se, but from the context that resulted when the subjects received different labels on different trials. Similar interference may have also arisen with the other 3 stimuli, but because performance on these stimuli was always imperfect, any such interference is not directly discerned. It is possible that 2 different processes, and perhaps 2 distinct mechanisms, underlie the effects of labeling, one process responsible for assimilation and the other for interference; alternatively, assimilation and interference might both result when a single mechanism that is geared to optimizing performance operates on different stimuli.

This interpretation of the results is reminiscent of a substantial number of findings from studies of selective attention in which subjects received multiple sources of stimulation and performance was assessed using paradigms developed by Garner (1974). By and large, these studies used measures of response time (RT) to gauge performance, often reporting evidence of 2 forms of cross-stimulus interaction, which can also be characterized as interference and assimilation. Melara and Marks (1990, Experiment 1) reported, for example, that subjects could identify the pitch of a tone (high vs. low) more quickly when the tone was accompanied by a corresponding visual bigram, either HI or LO, much as subjects in the present experiments more often identified a high-sucrose or high-citral flavor as mostly "sugar" or "citrus," respectively, when the flavors followed the corresponding label, SUGAR or CITRUS. In the experiment of Melara and Marks, performance was also poorer overall (RTs were longer) when the bigrams varied from trial to trial rather than remaining constant—evidence of interference, perhaps analogous to the results of the present Experiment 1. Although similar patterns have appeared in numerous other studies of selective attention (e.g., Marks 2004, Spence 2011), the mechanism that underlies interference and assimilation observed in tasks requiring rapid responding may differ from those observed in tasks that do not require rapid selective attention and discrimination.

Sensory versus decisional processes in assimilation

Although the evidence in the present study for label-induced interference is only modest, the evidence for assimilation of responses to the 2 labels is clear, and it raises the question: Through what kind of process or processes might subjects combine the linguistic and chemosensory information? The effects of labeling observed in both Experiments 1 and 2 can readily be described in terms of models that derive from signal-detection theory (Green and Swets 1966), as applied to taste and flavor (Linker et al. 1964; Irwin et al. 1992). A plausible, and simple, model would assume that linguistic labels affect flavor identification by modifying the locations of response criteria that are mapped onto the decision axis. By this token, we assume that each flavor stimulus in the present experiments produces, over many trials, a distribution of responses along an axis corresponding to a dimension of "sugar–citrus." Imperfect identification reflects the overlap between or among the distributions. In the absence of labels, and ignoring any source of interference, subjects likely set their criterion at or near the transition from sugar to citrus, responding "citrus" when the observation on a given trial falls below the criterion and "sugar" when the observation falls above it, as shown in the upper part of Figure 4. This figure illustrates a simple case with just 2 flavor stimuli, each of which produces a distribution of flavor responses. To simplify the exposition, the distributions in this example are equally displaced from the center of the axis.

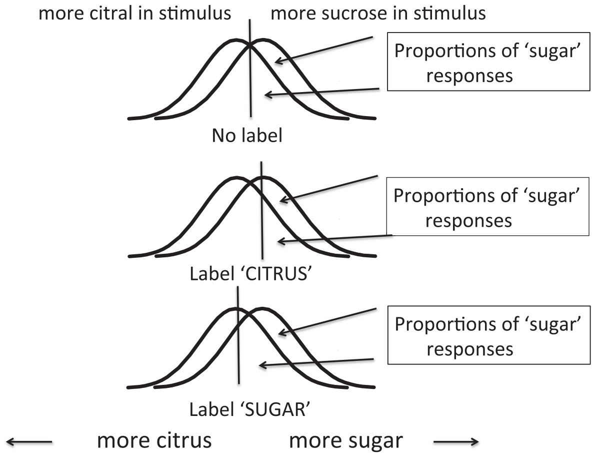

Figure 4 A signal-detection model to describe how labels could affect flavor identification by modifying the location of the response criterion. Each of the 3 identical pairs of curves shows hypothetical distributions of internal flavor responses along an axis representing "sugar–citrus." In each pair, the distributions on the left and right represent responses to a stimulus containing proportionally less and more sucrose, respectively. The area under each distribution to the left or right of each criterion corresponds to the probability of responding "citrus" and "sugar," respectively. Assuming that the criterion is set at a "neutral" location in the absence of any label (upper pair of distributions), then presenting the label CITRUS or SUGAR would shift criterion to the left (middle pair) or to the right (bottom pair), thereby correspondingly increasing the probability of responding "citrus" or "sugar."

According to this model, presenting a label, SUGAR or CITRUS, on a given trial shifts the criterion to the left or right, respectively. Assuming the sensory distributions are unchanged, shifting the criterion thereby leads to increases or decreases in the proportions of "sugar" and "citrus" responses (increasing or decreasing the area of each distribution that falls to the right of the criterion). With SUGAR, the criterion moves to the left, so the proportions of "sugar" responses increase in all flavor stimuli. With CITRUS, the criterion moves to the right, so the proportions of "sugar" responses decrease.

In a simple signal-detection model, where the underlying distributions of sensory events associated with the stimuli are Gaussian and have equal variance, the shifting-criterion model predicts a constant difference between effects of the labels CITRUS and SUGAR when the response proportions are transformed to normal deviates (z scores); because several of the individual proportions in the equaled 0.0 or 1.0, we followed Tukey's (1977) recommendation of first converting proportions into split scores (ss) by the formula ss = $(n + 1/6)/(N + 1/3)$ and then transforming ss into z scores. Reanalysis of the results of both Experiments 1 and 2 is compatible with this prediction: After transforming response proportions to z scores, the interaction between label and flavor stimulus was again not significant in either experiment: Experiment 1, $F(2,14) < 1$; Experiment 2, $F(3,51) = 1.38$, $P > 0.25$.

Finally, we mention an alternative interpretation of the model: It is possible that presenting a label does not only affect criterion but also (or instead) shifts the distributions of sensory responses themselves, the label SUGAR shifting the distributions toward the "sugar" pole and CITRUS shifting them toward the "citrus" pole while leaving the location of the criterion unchanged. The quantitative properties of models postulating shifting criterion and models proposing shifting sensory responses can be similar, even identical. More complex experimental paradigms than those used here would be needed to contrast these interpretations, which might apply also to analogous phenomena, such as the effects of color on flavor identification (Shavit AY, Marks LE, unpublished data).

One paradigm could capitalize on the use of RT as a surrogate for perceived flavor intensity (e.g., Yamamoto and Kawamura 1981; Bujas et al. 1989). Increases in stimulus intensity, hence in perceptual intensity, can lead to faster responses (smaller RTs) both in detection tasks, where subjects respond as quickly as possible to any stimulus, and in identification tasks, where subjects make different responses to different stimuli. When subjects rapidly identified low-frequency versus high-frequency tones with temporal uncertainty before each stimulus, greater sound intensity produced smaller RTs (Keuss and van der Molen 1982). Capitalizing on this finding, Arieh and Marks (2003) showed that when subjects rapidly identified low-frequency and high-frequency tones, raising the intensity context at either frequency led to greater RTs, as well as smaller judgments of loudness, but without the corresponding change in errors predicted by speed–accuracy trade-off. In this example, changing the stimulus context apparently changed the sensory representations of sound intensity without (just) shifting the response criteria.

Now back to flavor: If labels modify flavor identification by changing the underlying sensory representations (if SUGAR adds to perceived sucrose intensity and if CITRUS adds to perceived citral intensity), then it should be possible to mimic the effects of labeling by raising the concentration of sucrose or citral in each mixture to produce a comparable effect on identification. A critical test would then compare RT and errors across 3 conditions in a speeded flavor-identification task: a baseline condition presenting several sucrose–citral mixtures without labels; a labeling condition presenting the same flavor mixtures with labels; and an intensity-control condition presenting each mixture without a label but with an elevated concentration of sucrose or citral, predetermined to mimic the effect of labeling. Assuming that RTs are smaller in the intensity-control condition than in the baseline condition, without corresponding increases in errors, the RTs and errors observed in the labeling condition would be critical. If the addition of labels changes the sensory representations, then (as in the intensity-control condition), RTs should be smaller than RTs at baseline, without corresponding changes in errors. But if the labels simply lead to shifts in response criteria, then any decrease in RT should be offset by a corresponding increase in errors. Several methods are available for precisely determining the speed–accuracy relation for each stimulus (see, e.g., Arieh and Marks 2003, 2008).

Paradigms like the one just suggested, as well as the simpler paradigm of the present experiments, could prove useful when used in conjunction with statistical models (e.g., models based in Bayesian inference) that aim to predict optimal performance with multiples sources of information, such as those arising from sensory signals and verbal labels. Such paradigms might also shed light on the effects of labeling under a variety of sensory conditions, for example, with sensory disorders produced by disease or injury, and with sensory and cognitive aging.

Funding

This work was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health [R01 DC011823-01 and R01 DC009021-04 to L.E.M.].

Acknowledgments

We thank K. Botta for assistance with Experiment 2, and we thank International Flavors and Fragrances for providing the citral.

References

Arieh Y, Marks LE. 2003. Recalibrating the auditory system: a speed-accuracy analysis of intensity perception. J Exp Psychol Hum Percept Perform. 29:523–536.

Arieh Y, Marks LE. 2008. Cross-modal interaction between vision and hearing: a speed-accuracy analysis. Percept Psychophys. 70:412–421.

Auvray M, Spence C. 2008. The multisensory perception of flavor. Conscious Cogn. 17:1016–1031.

Bujas Z, Szabo S, Ajduković D, Mayer D. 1989. Individual gustatory reaction times to various groups of chemicals that provoke basic taste qualities. Percept Psychophys. 45:385–390.

Cerf-Ducastel B, Murphy C. 2001. fMRI activation in response to odorants orally delivered in aqueous solutions. Chem Senses. 26:625–637.

Cometto-Muniz JE, Cain WS. 1990. Thresholds for odor and nasal pungency. Physiol Behav. 48:719–725.

Distel H, Hudson R. 2001. Judgement of odor intensity is influenced by subjects' knowledge of the odor source. Chem Senses. 26:247–251.

Garner WR. 1974. The processing of information and structure. Hillsdale (NJ): Erlbaum.

Green BG, Dalton P, Cowart B, Shaffer G, Rankin K, Higgins J. 1996. Evaluating the 'Labeled Magnitude Scale' for measuring sensations of taste and smell. Chem Senses. 21:323–334.

Green DM, Swets JA. 1966. Signal detection theory and psychophysics. New York: Wiley.

Irwin RJ, Hautus MJ, Stillman JE. 1992. Use of the receiver operating characteristic in the study of taste perception. J Sens Stud. 7:291–314.

Kähkönen P, Tuorila H. 1998. Effect of reduced-fat information on expected and actual hedonic and sensory ratings of sausage. Appetite. 30:13–23.

Keuss PJ, van der Molen MW. 1982. Positive and negative effects of stimulus intensity in auditory reaction tasks: further studies on immediate arousal. Acta Psychol (Amst). 52:61–72.

Koza BJ, Cilmi A, Dolese M, Zellner DA. 2005. Color enhances orthonasal olfactory intensity and reduces retronasal olfactory intensity. Chem Senses. 30:643–649.

Lawless HT, Glatter S, Hohn C. 1991. Context-dependent changes in the perception of odor quality. Chem Sens. 6:349–360.

Linker E, Moore ME, Galanter E. 1964. Taste thresholds, detection models, and disparate results. J Exp Psychol. 67:59–66.

Marks LE. 2004. Cross-modal interactions in speeded classification. In: Calvert G, Spence C, Stein BE, editors. The handbook of multisensory processes. Cambridge (MA): MIT Press. p. 85–106.

Marks LE, Shepard TG, Burger K, Chakwin EM. 2012. Flavor-intensity perception: effects of stimulus context. Physiol Behav. 105:443–450.

Melara RD, Marks LE. 1990. Processes underlying dimensional interactions: correspondences between linguistic and nonlinguistic dimensions. Mem Cognit. 18:477–495.

Murphy C, Cain WS. 1980. Taste and olfaction: independence vs interaction. Physiol Behav. 24:601–605.

Schifferstein HN, Kole AP, Mojet J. 1999. Asymmetry in the disconfirmation of expectations for natural yogurt. Appetite. 32:307–329.

Small DM, Prescott J. 2005. Odor/taste integration and the perception of flavor. Exp Brain Res. 166:345–357.

Small DM, Voss J, Mak YE, Simmons KB, Parrish T, Gitelman D. 2004. Experience-dependent neural integration of taste and smell in the human brain. J Neurophysiol. 92:1892–1903.

Spence C. 2011. Crossmodal correspondences: a tutorial review. Atten Percept Psychophys. 73:971–995.

Tukey J. 1977. Exploratory data analysis. Reading (MA): Addison-Wesley.

Tuorila H, Cardello AV, Lesher LL. 1994. Antecedents and consequences of expectations related to fat-free and regular-fat foods. Appetite. 23:247–263.

Wansink B, van Ittersum K, Painter JE. 2005. How descriptive names bias sensory perception in restaurants. Food Qual Pref. 16:393–400.

Wilson RA, Mussinan CJ, Katz I, Sanderson A. 1973. Isolation and identification of some sulfur chemicals present in pressure-cooked beef. J Agric Food Chem. 21:873–876.

Yamamoto T, Kawamura Y. 1981. Gustatory reaction time in human adults. Physiol Behav. 26:715–719.

Yeomans MR, Chambers L, Blumenthal H, Blake A. 2008. The role of expectancy in sensory and hedonic evaluation: The case of smoked salmon ice-cream. Food Qual Pref. 19:565–573.