Detection of heteroquality taste mixtures

JOSEPH C. STEVENS

John B. Pierce Laboratory and Yale University, New Haven, Connecticut

Detection thresholds were measured for sweet (sucrose), salty (sodium chloride), sour (citric acid), and bitter (quinine hydrochloride) and for the 11 possible mixtures of these four substances. These 11 mixtures (6 binary, 4 ternary, and 1 quaternary) all turned out to be stimulus additive, in the sense that a person could reliably detect mixtures whose individual components are weaker than their unmixed thresholds. Tastants too weak to be perceived alone can thus make impact when in mixtures. The threshold concentration for a given compound was reduced in approximate proportion to the number of compounds added to it. This liberal heteroquality additivity contests the widespread belief that heteroquality mixtures (different chemicals evoking different qualities) are non-additive and homoquality mixtures (different chemicals evoking the same quality) are additive. Heteroquality additivity emerges on appropriate definition of the subject's task by forced choice (unavailable to earlier investigators), in order to skirt methodological pitfalls. Operating together, homo- and heteroquality additivity may concomitantly enable a person to sense natural mixtures of hosts of weak constituents, such as drinking water. In this regard, gustatory mixtures may function much as do mixtures of frequencies in audition and mixtures of gaseous compounds in olfaction.

This paper presents extensive evidence that a mixture of two or more stimuli having different taste qualities—say, salt and sugar—may be reliably detectable even though its components are too weak to be reliably detected alone. Supporting data are given for 11 heteroquality (also called cross-quality) mixtures of four compounds representing four qualities: sweet (as evoked by sucrose), sour (by citric acid), saltiness (by sodium chloride), and bitter (by quinine hydrochloride). The 11 mixtures comprise the 6 binary ones, the 4 ternary ones, and the 1 quaternary.

This research differs from Hahn and Ulbrich's (1948) often cited study and from scattered subsequent reports primarily in (1) use of whole-mouth, sip-and-spit sampling, (2) criterion-free psychophysical measurement, and (3) vastly greater emergence of additivity in heteroquality mixtures. Since Hahn and Ulbrich's study, it has been widely accepted that homoquality mixtures are additive but that heteroquality mixtures are usually nonadditive and only occasionally additive.

It is crucial to distinguish between the detection of mixtures themselves, which, as should become clear, appear to be "additive" by nature, and the detection of a particular component in a mixture. The latter kind may be called "subtractive" or "inhibitory," because the threshold for one of the components is raised by the presence of other compound(s); for example, it is harder to detect sugar in lemonade than in water, because the sugar is "masked"

by the other (mainly sour) constituents of the lemonade. In signal detection terms, the subject's task for additive mixtures is to decide between a sample containing the mixture and a sample containing deionized water only. In contrast, for subtractive mixtures the subject must decide which of two samples contains a given compound—say, sugar—in mixture with another—say, sour—compound (or compounds) as opposed to the other compound (or compounds) alone. Failure to make this distinction consistently may be a key (but not the only) reason why the literature on taste mixtures has been judged as equivocal (see, e.g., Bartoshuk, 1975, 1988; Geldard, 1972; McBride & Johnson, 1987; Moskowitz, 1972; Pangborn & Trabue, 1967; Pfaffmann, 1971). Although this paper deals with additive mixture, subtractive mixture has also come under study and is in progress. Results have already been reported on mixtures of sucrose, citric acid, and sodium chloride. These show that the threshold for detecting any one of them increases with increasing concentration of the other in mixture (Stevens & Cain, 1993; Stevens, Cain, Demarque, & Ruthruff, 1991), to as much as ten-fold.

Thresholds in subtractive mixtures are similar but not quite identical to the so-called recognition threshold, which is the minimum stimulus level necessary to identify the specific quality or type of a sensation, usually higher than the detection threshold (Pfaffmann, Bartoshuk, & McBurney, 1971). When a subject has to decide between a mixture of sour and sweet and a sweet alone, as in the example above, qualitative recognition may be implicit but is not necessarily specifically called for.

Actually, definition of the task in terms of forced choice is definitive and unambiguous for understanding the difference between additive and subtractive mixture. Both are useful because they can map in profoundly dif-

This research was supported by NIH Grant AG 04287. The author is grateful to Julianne M. Hoffman for extensive technical assistance and to Lawrence E. Marks, Enrique Cometto-Muñiz, William S. Cain, Janneane F. Gent, and Matthew Q. Patterson for technical and editorial advice. Correspondence should be addressed to J. C. Stevens, John B. Pierce Laboratory, 290 Congress Ave., New Haven, CT 06519.

ferent ways key aspects of everyday perception. These operational differences in task, however, fall short of describing the subjective experiences in discrimination. The latter can only roughly be characterized from introspective reports. A subtractive mixture tends to be quality selective; that is, one distinguishes between "something" (sucrose plus NaCl, say) and "something else" (NaCl alone, say). Both the signal stimulus (sucrose plus NaCl) and the blank (NaCl) usually have obvious, even strong taste; they differ only with respect to the presence of sucrose, whose threshold in the mixture is sought and compared with its threshold unmixed. In contrast, an additive mixture tends to be quality blind; one distinguishes between "something" (sucrose and NaCl, say) and "nothing" (water), and the "something" may occasion one quality or the other quality, both of them, or neither—for example, it could be an emergent (synthetic) quality in the manner suggested by Erickson and Covey (1980). In fact, subjects report incidentally experiencing different qualities, always weak, only some of the time. In testing additive mixtures, in contrast to subtractive ones, the stimuli, both signals and blanks, are always weak in taste. Thus, the subjective nature of the discrimination resists rigorous defining. Yet it is nonetheless clear that mixing permits us to detect substances that are separately undetectable.

It may help to understand the rationale here to explain that, like the earlier threshold studies of gustatory, olfactory, and auditory mixtures discussed later on, this one defines additivity in stimulus terms. (In contrast, and often causing confusion, suprathreshold studies of mixtures, as will be discussed later, have usually been described in response [subjective] terms, such as magnitude estimates.) In the present study, additivity is said to obtain unequivocally if and only if the concentration of each component of the just-detectable mixture is weaker than its just-detectable concentration tasted alone. Defined thus, the mixtures studied turned out to be pervasively additive. The apparent degree of additivity ranged from mildly hypo-additive to simple additive to mildly hyperadditive. Simple (complete) additivity means that the concentration of each component in mixture equals $1/n$ times its concentration measured separately, where n equals 2 for binary mixture, 3 for ternary mixture, and 4 for quaternary mixture. Hypoadditivity means that the components averaged somewhat larger than $1/n$, and hyperadditivity, the more frequent outcome, that the components averaged somewhat less than $1/n$.

The basic plan of the study was (1) to make up mixtures whose constituent compounds had concentrations proportional to thresholds measured separately in earlier studies, (2) in a given experimental session to measure, by forced-choice up-down tracking, thresholds for one given mixture and each of its component compounds, and (3) to carry out this measurement for all 11 mixtures in each of 11 subjects, yielding 429 thresholds altogether.

METHOD

Subjects

Five men and 6 women between 18 and 27 years of age took part in 11 test sessions each, on separate days spread over a period of approximately 2 months. In each session, the subject gave a threshold for 1 of the 11 mixtures plus one each for each of the components of that mixture. This meant that three, four, or five thresholds were measured, depending on whether the mixture was binary, ternary, or quaternary, and the duration of the session varied accordingly—from about 45 to 120 min. The order in which the various mixtures came under study was varied greatly from one subject to another to avoid possible order effects. The subjects gave informed consent and received a payment of \$150.

Stimulus Materials

Eleven solution series, each comprising 15 concentration levels, were made up by liquid dilution with deionized water (DHOH) of Baker-grade sucrose (S), citric acid (C), sodium chloride (N), and reagent-grade quinine hydrochloride (Q), as well as by mixtures of these. The solutions were stored under refrigeration, warmed as needed to room temperature before testing, and discarded after 2 weeks.

To ensure that the mixtures contained reasonable proportions of the constituent compounds (i.e., that all constituents would contribute approximately equally), it was necessary to estimate in advance the detection threshold for the four compounds separately. The estimates were the mean thresholds from young subjects in a concurrent project on subtractive mixtures mentioned earlier and reported in part by Stevens et al. (1991) and Stevens and Cain (1993); these were based on 30 subjects' thresholds for S and C, 15 subjects' thresholds for N, and 4 subjects' thresholds for Q. These means are listed in Table 1, together with the means of the corresponding thresholds obtained in the present study (for each compound, 77 thresholds—7 from each of 11 subjects) and averages obtained by Bartoshuk, Rifkin, Marks, and Bars (1986) by using a very similar method (as explained by Bartoshuk, 1975) on 18 young subjects. All three studies are in first-order agreement; the reasonably good agreement between preselected values and subsequently measured ones implies that the mixture concentrations were favorable, though the measured thresholds for S did fall slightly above preselected values, whereas those for C, N, and Q fell about the same amount below preselected ones. This means that the proportions of S to the other components of the mixtures may have been, on the average, a little too low to be precisely representative of the relative sensitivities of the present subjects. Just how this slight imprecision manifests itself in the outcome will become clear later.

Each series of 15 concentrations was constructed by successive 0.25 log dilutions starting from mother solutions (0 dilution step) of 0.399 mole of S, 0.0042 mole of C, 0.12787 mole of N, and 0.00025 mole of Q in 1,000 ml of solution. This arrangement placed the best estimation of the threshold, as given in Table 1, at the eighth dilution step (-2.0 log units or $1/100$ below the mother

Table 1
Mean Threshold Concentration for Sucrose (S), Citric Acid (C), Sodium Chloride (N), and Quinine Hydrochloride (Q) in Three Studies

Compound	$\times$ Concentration	Bartoshuk et al., 1986	Concurrent Project	Present Study
S	10^{-3}	2	4.0	5.3
C	10^{-5}	2	4.3	3.2
N	10^{-4}	9	12.8	8.9
Q	10^{-6}	1	2.5	1.6

Note—The values in columns 3, 4, and 5 multiplied by the values in column 2 give molar concentration.

Figure 1. Detection thresholds (means and standard errors) for six binary mixtures (left), four ternary mixtures (middle), and one quaternary mixture (far right) of four compounds. The position (along the ordinate) of each plotted letter stands for the ratio (times 100 to convert to a percentage) of the threshold concentration of a compound in mixture to the corresponding threshold concentration unmixed. The filled circles represent the geometric average of all such percentages associated with a given mixture. The smaller the percentage, the greater the degree of additivity. The dotted line segments mark the boundary between hyperadditivity and hypoadditivity for the three types of mixture: binary (50%), ternary (33%), and quaternary (25%).

concentration), at the middle of each 15-step series. Mixture series were made up in the same way, except that the mother solution (1,000 ml) contained the above-mentioned number of moles of more than one substance—for example, 0.399 mole of S and 0.0042 mole of C in the 1,000-ml mixture of S and C. In such a mixture the eighth dilution step contained the same amounts of S and C as in the unmixed series. Thus, if the mixtures are additive, the threshold concentration should fall at a dilution step greater than eight (i.e., both concentrations lower than at their eighth step). This was the typical result.

Psychophysical Method

The method was a forced-choice version of up-down tracking (Wetherill & Levitt, 1965). In any given session a subject generated separate, independent threshold tracks for the mixture and for each of its components, but these tracks were closely intertwined by alternating tracks on each trial. For example, if the mixture was SC, the order of the stimuli might be S, C, SC, S, C, SC, S, . . . to the very end of the track. For another subject, the order might be SC, C, S, to avoid potential order effects. The rationale for intertwining tracks is that fluctuation or drift of the subject's sensitivity or criterion over time, as a result of practice or fatigue, say, would tend to influence all tracks of a session equally, and that the optimum measure of additivity might be achieved by placing mixture and components in proximity. However, to discount the possibility that intertwining the tracks might be a *necessary* condition for demonstrating additivity, 8 of the original 11 and 3 new sub-

jects took part in an additional, ancillary session testing S, C, and SC in three successive and counterbalanced, rather than intertwined, tracks.

On each trial, the subject had to decide between two 30-ml plastic medicine cups, one of which contained about 5 ml of DHOH only (the blank), the other, about 5 ml of a mixture or one of its components (the signal), as to which had a taste or which had the stronger taste. The signal and the blank were placed simultaneously in front of the subject in random left-right position from trial to trial. The sip-and-spit method was used; at the start of the session, and after each sampling, the subject rinsed thoroughly with DHOH. Solutions, both signals and blanks, were swirled around in the mouth briefly and expectorated.

The tracks were generated according to the following rules: (1) The starting concentration for every track was the eighth dilution step. This means that a track started with the best estimate of the threshold for a component, but, presuming a degree of threshold additivity, somewhat above the threshold for a mixture. This rule was chosen to avoid bias favoring additivity, but in so doing possibly introduced a small bias in the opposite direction. (2) No data were recorded for a given track until the subject erred on a trial; this rule, recommended by Wetherill and Levitt, tends to reduce bias associated with starting the track at a level too far removed from the threshold; as explained later, this rule presumably helped, but it did not entirely eliminate the bias. (3) Whenever a subject erred on a trial, the concentration on the next trial of that track was increased by one step (i.e., to the next higher dilution

step); whenever a subject chose correctly on a trial, the same concentration was presented on the next trial; whenever the subject chose correctly on two successive presentations of the same concentration on the same track, the concentration was decreased by one step on the next trial of that track. This rule is often called the "one incorrect-up, two correct-down" rule, and on the average it results in (estimated) 70.7% (i.e., $100 \times 0.5^{0.5}$) correct responding. (4) Tracks continued in alternation this way until seven or more transitions had occurred in the direction of the track from higher to lower or from lower to higher concentration ("reversals"). Testing ceased when each track exhibited at least seven reversals.

Method of Calculating Threshold

As is customary, threshold was estimated from the transition levels of the track. The first transition level was never counted; the remaining transition levels (dilution steps), labeled reversals one through six, were counted, but not reversals beyond these six that occurred when a track lasted beyond the original seven; thus all estimates of the threshold were, for uniformity, based on six levels only, even though a track might have gone on longer. The question arises, what is the best estimate of the threshold based on these six levels? To begin with, we tried the mean of all six of them, the mean of the last four, and the mean of the last two. All three of these yielded a clear and basically similar pattern of additivity, but greater reliance on the tail end of the track suggested somewhat greater degree of additivity. Why should this be?

The reason why emerged from plots of the subjects' mean transition level as a function of transitions one through six. The average track for the quaternary mixture slopes away from the starting level (the eighth dilution step) more so than the average tracks for the ternary mixtures, which in turn slope away more than those for the binary ones, which in turn slope away more than those for the individual compounds. This is the inevitable consequence of the decision to begin all tracks at the same level. For example, the final results, as we shall see, suggested that additivity for the quaternary mixture was approximately "complete," or that its constituent concentrations should equal about one fourth their individual, separate concentrations. This implies that the quaternary track would have to shift upward by about 2.4 dilution steps—that is, $(\log 4)/0.25$ —from its starting point to the sixth transition level. It would also imply that including early stages of the tracks in the averages would guarantee an underestimation of the degree of additivity.

A solution that took into consideration all six transition levels but also reckoned with the sloping nature of some of the tracks was (for each track) to fit a regression line to the transitions levels, y , as a function of the transition number, x , from $x = 1$ to $x = 6$. Threshold was reckoned as the calculated level of y at $x = 6$. This

value has the advantages that (1) it best reflects the status of the track near its end, where it has had the chance to approach its asymptotic level, and (2) it is stabler than the actual sixth step itself, which, for one thing, was always a downward transition. The solutions obtained by this procedure approximate simple averages of the fifth and sixth transitions, but are somewhat less arbitrary and a little more realistic, and are thus offered as satisfactory representations of the data.

RESULTS

As mentioned, additivity is most convincing when in a given test session the threshold of every component in a mixture turns out to be lower than its threshold unmixed. The a priori probability of this happening by chance is 1/4 for a binary session, 1/8 for a ternary one, and 1/16 for a quaternary one. In the binary sessions, this stringent criterion was met 45 times out of 66 (computed by binomial expansion, throwing out occasional ties; overall $p < 10^{-8}$); in the ternary sessions, 31 times out of 44 (overall $p < 10^{-8}$); and in the quaternary sessions, 9 out of 11 times (overall $p < 2 \times 10^{-5}$). The ancillary study of S and C mixture, using successive, nonintertwined tracks, also yielded reliable additivity, in 8 of 11 measurements in which mixture thresholds fell below both component thresholds (overall $p < .009$). These simple nonparametric statistics speak to the pervasive character of heteroquality additivity.

The parametric results appear in Table 2 and Figure 1. Table 2 lists the average¹ molar concentration of each component at the threshold of each of the 11 mixtures (right side) and the average molar concentration of the corresponding unmixed component (left side). In every case, the unmixed component exceeded the corresponding mixed component. By "corresponding concentrations" is meant concentrations of the relevant compound obtained in intertwined tracks obtained in the same test session. Figure 1 was derived from the data in Table 2 to reveal several telling features. Here is plotted the ratio of the concentration of each mixed component to its corresponding unmixed concentration, expressed as a percentage, for all the mixtures. These percentages are al-

Table 2
Projected Molar Concentrations of the Mixture Constituents in Mixture (Right) and Unmixed (Left)
Sucrose (S), Citric Acid (C), Sodium Chloride (N), and Quinine Hydrochloride (Q)

Mixture	Concentration							
	Unmixed				Mixed			
	[S] $\times 10^{-3}$	[C] $\times 10^{-5}$	[N] $\times 10^{-4}$	[Q] $\times 10^{-6}$	[S] $\times 10^{-3}$	[C] $\times 10^{-5}$	[N] $\times 10^{-4}$	[Q] $\times 10^{-6}$
SC	6.93	4.00			1.50	1.61		
SN	4.40		10.8		1.26		4.03	
SQ	5.10			2.35	1.43			0.896
CN		3.63	11.1			1.63	4.87	
CQ		3.64		1.23		1.52		0.891
NQ			8.39	1.97			5.59	1.09
SCN	6.51	4.19	12.2		1.35	1.44	4.32	
SCQ	3.79	2.26		1.20	0.589	0.640		0.375
SNQ	5.37		5.80	1.33	0.668		2.14	0.418
CNQ		1.85	7.82	1.67		1.15	3.44	0.672
SCNQ	5.74	3.93	7.82	2.14	0.886	0.948	2.84	0.555

ways smaller than 100. Their averages (geometric means) for each mixture are also plotted, as filled circles, in the figure. The dotted line segments mark the boundary between hypoadditivity and hyperadditivity. This boundary is 50% for binary, 33% for ternary, and 25% for quaternary mixtures.

Since all the circles fall below 100%, all the mixtures were additive to some degree. A few lie above the dotted boundary and appear therefore to be hypoadditive; more lie below it and appear therefore to be hyperadditive. That mixtures were additive was further substantiated statistically by paired *t* tests on the 39 percentages shown in Figure 1, both for each component of a mixture (28 letters) and the averages (11 circles) of the components of a mixture. The null hypothesis (that the population percentage equals 100) was rejected at $p < .05$ in 37 of 39 cases (typically at a low or very low level of p), the non-significant exceptions being $p < .12$ for the component Q in the mixture CQ and $p < .15$ for the component N in the mixture NQ. All 11 averages (circles) were significant at $p < .02$ to $< 10^{-5}$. Each of the 11 subjects showed reliable additivity when his/her data were examined across all 11 mixtures, at $p < .05$ to $p < 10^{-4}$ (mdn $p < 10^{-3}$). Lower individual reliability appeared to stem more from day-to-day variation than from overall degree of additivity. Finally, no significant gender differences emerged.

That the plotted values for sucrose in Figure 1 (S symbols) all fall relatively low in the figure reflects the fact that the proportion of sucrose to the other components of the mixture was a little lower than anticipated, as ex-

plained in the Method section. There is a hint, but no statistical evidence, that sucrose played a more potent role in mixtures than did the other three tastants. Analysis of variance (operating on the *circles* in Figure 1) comparing the three binary mixtures containing sucrose with the three not containing it was not significant, nor was an analysis of variance comparing the three ternary mixtures containing sucrose with the one ternary mixture not containing it.

This judgment about sucrose's role in mixture rests on the *circles* in Figure 1, not, as it might superficially appear, on the plotted values of S, which fall relatively low in the figure; they do so because the proportion of sucrose to the other components of the mixture was a little lower than the proportion of sucrose threshold to the other thresholds measured in the present experiment. It is therefore irrelevant, and potentially misleading to the unwary, that the value of S is relatively low in all seven mixtures of which it was a component.

Another feature of mixture, difficult to appreciate from Figure 1, reveals itself in the log-log coordinates of Figure 2. Here is plotted again the concentration of a mixed component as a percentage of its unmixed concentration, in binary, ternary, and quaternary mixture. The geometric averages of these percentages across compounds are also plotted (as filled circles). The line in Figure 2 has a slope of -1.0 (computed by least squares, -0.994). This line would signify that the threshold concentration goes down in exact proportion to the number of components to which it is added, or, in other words, where Y is threshold concentration (expressed as a percentage) and X is number of compounds,

$$\log Y = -\log X + \log k, \text{ or } YX = k. \quad (1)$$

As can be seen, especially for the circles, this equation describes the outcome to a first order of approximation. This is the same simple reciprocity that describes so many summation phenomena in sensory psychophysics in general.

As stated, this generalization may gloss over possible real differences in degree of additivity of particular mixtures. Although additivity may indeed be the general rule, it cannot take place to the degree that one compound might suppress the activity of another at the peripheral (e.g., receptor) level. After all, what fails to get through the periphery can hardly be expected to add to other events further upstream. By means of split-tongue stimulation and magnitude estimation, Kroeze and Bartoshuk (1985) claimed, for example, to demonstrate peripheral partial suppression by NaCl of the bitterness of quinine hydrochloride. (See also Schifferstein & Frijters, 1992.) Partial suppression of this kind might help to explain, for example, why the binary mixture NQ lies well above the dotted line in Figure 1 and fell short of statistical significance. On the other hand, it is also important to note from Figure 2 that, for each of the four substances, the concentration falls monotonically as the number of components increases from one to four. In

Figure 2. Threshold reduction (concentration in mixture divided by concentration unmixed times 100) as a function of number of compounds mixed. Letters stand for individual compounds; filled circles, for geometric average percentages for the component compounds. To the extent that these data fit a line with negative unit slope in log-log coordinates, they imply complete additivity regardless of the number of components.

other words, summation takes place freely with increasing complexity of mixture, and there is, in fact, no evidence that, on the average, degree of additivity is any smaller in quaternary mixtures, say, than in binary ones.

DISCUSSION

Additivity of Heteroquality Mixtures

As stated earlier, homoquality taste mixtures are often presumed to be additive. Evidence comes from (1) supra-threshold mixture studies—the sum of a subject's taste estimates of the components in the mixture often equals, or occasionally even exceeds, the sum of the estimates of the components tasted separately. (Note that here additivity is couched in subjective rather than stimulus terms.) Evidence comes also (2) from Hahn and Ulbrich's (1948) prodigious threshold study of scores of binary mixtures, of all qualities, with additivity defined here in stimulus terms. It turned out that concentrations a and b of two homoquality, mixed compounds can be traded reciprocally for each other to preserve threshold, or, $ab = k$. This is a stringent definition of additivity, implying, for example, that if a is doubled b can be halved, and these authors reported it to be the rule.

Results for heteroquality mixtures were, unfortunately, less clear-cut, occasionally favoring additivity, but most other times not. The reason may have to do with the exact formulation and understanding of the subject's task, in the absence of forced-choice methodology. For homoquality mixtures, the task is relatively straightforward: the subject reports either sensing or failing to sense the mixture, or any of its components. Whether the subject senses the mixture or a component, the sensation has the same quality. (This may not always be *strictly* true because at least some of Hahn and Ulbrich's substances may have had a secondary qualitative component, as Bartoshuk [1988] pointed out in her evaluation of Hahn's findings on cross-adaptation of same-quality tastes.) Overall, though, qualitative uniformity simplifies the measurement of additivity.

In contrast, heteroquality mixtures potentially suffer ambiguity. One can from time to time pick out the different qualities in the mixture. If the subject opts systematically to respond positively to only one or another of the qualitative components of the mixture, the resulting threshold can actually be raised rather than lowered by the other components. Without rules to define the tasks unambiguously, "detection threshold" could turn out to be additive or subtractive. Fortunately the invention of forced-choice methodology makes it possible to avoid ambiguity. But classical psychophysical methods, like constant stimuli without forced choice, could muddy the distinction. This may help to explain why heteroquality additivity turns up ubiquitously in the present study but only erratically heretofore.

Other factors, though, besides forced-choice methodology could also contribute to the dissimilar pictures of heteroquality additivity furnished by the present study

and by Hahn and Ulbrich's (1948). Most noteworthy, after psychophysical task, was mode of stimulus delivery. Instead of whole-mouth, sip-and-spit sampling, Hahn and Ulbrich flowed their solutions across a portion of the tip of the tongue. That the difference between these two delivery modes might matter significantly, even crucially, to the rules of heteroquality mixture was already argued by Bartoshuk (1975). Yet even though Hahn and Ulbrich's procedure would seem to be uncharacteristic of the normal way of perceiving tastes, it seems improbable that stimulus delivery alone could account for the differences reported between the two kinds of mixture, given that the same delivery system was used for measurement of both. Furthermore, one must underscore the point made earlier that a particular mixture, binary ones especially, may exhibit less or even no additivity because of peripheral suppression. Similar threshold measurements of other pairs of compounds, including various Na bitter ones, would seem to be called for.

Threshold Versus Suprathreshold Mixtures

Since Hahn and Ulbrich (1948), suprathreshold research on mixtures has held center stage. A complete account and evaluation would go far beyond the scope of this paper, but a very brief history may be helpful. The introduction of the method of magnitude estimation by S. S. Stevens (1957) and its modification by Smith and McBurney (1969) to include multiquality estimations of taste stimuli doubtless helped to foster suprathreshold investigation. Studies by Moskowitz (1972) and Bartoshuk (1975) set the stage by obtaining magnitude estimates for each of the various qualities in a suprathreshold mixture. In Bartoshuk's study, the stimuli were binary, ternary, and quaternary mixtures of sucrose, NaCl, hydrochloric acid, and quinine hydrochloride, clearly a heteroquality set much as in the present threshold study. Many studies have followed (for empirical and theoretical accounts see, e.g., Frijters, 1987; Frijters & De Graaf, 1989; Kroeze, 1989; McBride, 1989; Schifferstein & Frijters, 1993). In studies that, like Bartoshuk's, have employed heteroquality mixtures, the magnitude estimations of the various component qualities in the mixture turn out to be smaller than those of the same components sampled alone. In this sense, heteroquality suprathreshold taste mixtures have been characterized as mutually suppressive (less than fully additive). In contrast, in similar homoquality studies, the estimation of mixture's magnitude sometimes falls short of but more often equals or even exceeds the sum of the estimations of the same components sampled alone. In this sense, homoquality mixtures have been said to be more additive (even synergistic) than heteroquality ones.

Applied to heteroquality mixtures, but probably not homoquality ones, the magnitude estimation procedure, focusing the attention analytically as it does on the individual components, may favor a result that is more analogous to subtractive threshold measures than to additive ones. (To the extent that homoquality mixtures are truly

homoquality, it would be impossible for the subject to differentiate their components.) A few taste scientists (e.g., McBride, 1989; Schifferstein & Frijters, 1993) have, however, given recognition to the overall "impact" of these mixtures: on their view, the impact of a mixture may not necessarily equal the simple algebraic sum of its components; sometimes a person may be more concerned with whether a taste is weak or strong overall than with its qualitative makeup. "Impact" in this sense may be an appropriate analogue to the additive mixtures measured here, for which qualitative makeup is also incidental to detection. In psychoacoustics, distinction is made between the overall loudness of a tonal complex and the loudness of its frequency components; both can be measured by matching with a standard probe sound (without, in fact, necessarily resorting to magnitude estimation). Both kinds of loudness have interest in everyday life—for example, one's concern may be the overall loudness of a radio or the relative loudness of its bass component. The literature on auditory mixtures is voluminous; suffice it to say that every taste phenomenon discussed here has a relevant mixture analogue there (see, e.g., Green, 1976; Luce, 1993).

The operations for measuring threshold and suprathreshold mixtures and the terminology adopted for describing the results makes formidable the search for a unified theory. Already cited is the practice of using stimulus versus response terms for characterizing additivity, suppression, and so forth. The theoretical accounts listed earlier list alternative models of suprathreshold mixtures that still need to be reconciled with each other, let alone with threshold data. It is hoped, though, that a step in the right direction may be to clarify the nature of mixture thresholds, both additive and subtractive, and homo- and heteroquality.

Models for Heteroquality Threshold Additivity

In the long run, then, the facts concerning mixture threshold need to be integrated into more general models such as those formulated by McBride (1989) and Schifferstein and Frijters (1993). When it comes to thresholds per se, however—especially additive ones—two general approaches are possible. These may be called common stimulus path and probabilistic models. By the former is meant that the neural signals for the various taste qualities may feed at some stage along the neural path into a common detection or "impact" channel, while at the same time preserving their independent status in parallel channel(s) for quality. Such a model places emphasis on stimulus additivity or nonselectivity, but complete stimulus additivity across quality is not necessary to the model since the inputs from the various qualities could be weighted; for example, sweet could contribute more heavily to the common channel than do the other qualities. Further testing of stimulus additivity in heteroquality mixtures might do well to include tests of Hahn and Ulbrich's (1948) reciprocity function, which describes the results of their binary homoquality mixtures.

By *probabilistic model* is meant that the psychometric function for a mixture might be predictable from the psychometric functions for the separate components. According to this model, the probability of detecting a mixture would equal the sum of the probabilities of detecting each of the components separately (subtracting the probability of detecting the components jointly). If mixtures work in this way, then the threshold, defined as the level of the mixture detected correctly—say, 70% of the time, as in the present study—would in fact be lower than the thresholds of the components.

This is a testable model, but to test it would require the time-consuming measurement of the separate psychometric functions for each component and for the mixture, presumably by combining the method of constant stimuli with forced choice. To do this for all of the qualities and their various mixtures would pose a formidable undertaking, but would perhaps be worthwhile if only for an exemplary mixture or two to determine how well the model works.

The notion of parallel processing of the different qualities and overall intensity is, of course, quite speculative in taste. In vision, the idea of parallel processing is hardly a stranger. Separate parallel channels for processing such properties as color, lightness, form, and space have received widespread acceptance as one of the key features of visual anatomy and physiology (Zeki, 1993). An advantage of this model is that it can be made to apply to both threshold and suprathreshold additivity; in contrast, it is hard to imagine how a probabilistic model could be extended to suprathreshold additivity.

Until one of these two types of model or some other prevails, it will have to suffice to describe threshold additivity in stimulus terms and in those terms to try to grasp what meaning, if any, additivity of heteroquality mixtures has for understanding the everyday world of taste, a question already addressed with some insight in psychoacoustics and olfaction.

Taste Mixtures in Everyday Life

To appreciate the biological function of mixture additivity in taste, it may be useful to ponder the other senses. In vision, threshold and suprathreshold mixtures of different wavelengths have long been known to be basically luminance additive (Abney's law). Perhaps in better analogy to taste, additive and subtractive mixtures have come under study in both olfaction and hearing. In olfaction, subtractive mixtures have predominated, usually under the rubric of masking or counter-odorization (Cain & Drexler, 1974). An important exception was Guadagni, Buttery, Okano, and Burr's (1963) early study of the additivity of olfactory mixtures. Their study was prompted by the uncovering, via the then new technology of gas chromatography, of the greatly complex olfactory makeup of fruits and vegetables, which may include hundreds of organic compounds, presumably individually undetectable, but perceptible in concert. These authors believed that complete additivity might be limited to members within

a chemical class (e.g., the saturated aldehydes), but they conjectured that partial (hypo-) additivity might extend across classes; actually, later work suggests that full or even hyperadditivity might extend even across chemical classes (Cometto-Muñiz & Cain, 1994; Laska & Hudson, 1991; Patterson, Stevens, Cain, & Cometto-Muñiz, 1993), pushing the "gain" on sensitivity as a result of mixture additivity beyond known limit. Without such additivity, many common objects might otherwise make no impact. For example, one might fail to distinguish fresh air from stale air containing a host of subthreshold contaminants.

The psychoacoustic world works much the same. Its literature encompasses knowledge of how the threshold for virtually any acoustic signal depends on the spectral makeup and strength of almost any conceivable context in which it may occur (subtractive type), to the measurement of the threshold of a complex of tones as a function of the number and spacing of its tonal components (additive type). Both kinds of threshold have their place. Mixtures of frequencies (in sufficient complexity, noises) make their impact when individual tonal components are subliminal. This fact is implicit in acoustical psychophysics, but its relevance to olfactory and gustatory mixtures can perhaps be appreciated most dramatically in the work of Gässler (1954). He demonstrated that complete stimulus (energy) summation can take place over at least 10 different tones spaced within a critical band. Hypoadditivity took place over at least an additional 40 or so frequencies beyond the critical band. Thus the ear exhibits tremendous power of energy summation across frequency. At the same time, it exhibits exquisite analytic capacity when a person is directed to respond to a particular kind of signal as opposed to the acoustical context in which it is embedded. In other words, auditory thresholds, like gustatory ones, can be additive or subtractive, and the psychophysical rules for these two kinds of mixtures are quite different.

Heteroquality additivity revealed here and homoquality additivity revealed earlier (Hahn & Ulbrich, 1948), acting concomitantly, could extend additivity over tens or even hundreds of compounds. Limits on such integration, if they exist, are unknown. Although Hahn and Ulbrich never went beyond binary mixtures, there is no reason to suppose that additivity stops there. It remains for future research to determine, first, whether other tastants besides the four studied here also show such generous heteroquality additivity, and second, how far mixture can operate to increase the "gain" on gustatory sensitivity. In theory, mixture additivity is by no means limited to just two, three, or four components. It is possible, for example, that a person can distinguish between plain water and drinking water containing a host of minuscule amounts of various solutes by themselves far subliminal, just as the nose can distinguish fresh from stale air. It thus remains to be worked out how and to what extent additive mixture operates in everyday life, but its role can hardly be considered minor as Hahn and Ulbrich seemed to suppose.

Perusal of the psychophysical data on mixtures in the various senses thus reveals their immense relevance to understanding the work of the senses. In the desire to simplify, the sensory scientist seeks first to reduce stimulation to its simplest elements—single wavelengths for light, single tones for hearing, single gaseous solutions for smell, and single aqueous solutions for taste. Initially, the complex stimuli of everyday life are typically much too difficult to deal with if order is to emerge. Eventually, however, we must face the fact that most objects of perception are made up of a myriad of component stimuli and that we need to know how the senses deal with them in concert. Taste is no exception.

REFERENCES

- BARTOSHUK, L. M. (1975). Taste mixtures: Is mixture suppression related to compression? *Physiology & Behavior*, **14**, 643-649.
- BARTOSHUK, L. M. (1978). The psychophysics of taste. *American Journal of Clinical Nutrition*, **31**, 1068-1077.
- BARTOSHUK, L. M. (1988). Taste. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' Handbook of experimental psychology* (2nd ed., pp. 461-499). New York: Wiley.
- BARTOSHUK, L. M., RIFKIN, B., MARKS, L. E., & BARS, P. (1986). Taste and aging. *Journal of Gerontology*, **41**, 51-57.
- CAIN, W. S., & DREXLER, M. (1974). Scope and evaluation of odor counteraction and masking. In W. S. Cain (Ed.), *Odors: Evaluation, utilization, and control* (Annals of the New York Academy of Sciences, Vol. 237, pp. 427-439). New York: New York Academy of Sciences.
- COMETTO-MUÑIZ, J. E., & CAIN, W. S. (1994). *Sensory additivity of airborne chemicals in mixtures to elicit thresholds for odor, nasal pungency, and eye irritation*. Unpublished manuscript.
- ERICKSON, R. P., & COVEY, E. (1980). On the singularity of taste sensations: What is a taste primary? *Physiology & Behavior*, **25**, 527-533.
- FRIJTERS, J. E. R. (1987). Psychophysical models for mixtures of tastants and mixtures of odorants. In S. D. Roper & J. Atema (Eds.), *Olfaction and taste IX* (pp. 67-78). New York: New York Academy of Sciences.
- FRIJTERS, J. E. R., & DE GRAAF, C. (1989). Modeling taste mixture interactions in equiratio mixtures. In D. G. Laing, W. S. Cain, R. L. McBride, & B. W. Ache (Eds.), *Perception of complex smells and tastes* (pp. 245-264). San Diego, CA: Academic Press.
- GÄSSLER, G. (1954). Über die Hörschwelle für Schallereignisse mit verschieden breitem Frequenzspektrum, *Acustica*, **4**, 1954.
- GELDARD, F. A. (1972). *The Human Senses*. New York: Wiley.
- GREEN, D. M. (1976). *An introduction to hearing*. Hillsdale, NJ: Erlbaum.
- GUADAGNI, D. G., BUTTERY, R. G., OKANO, S., & BURR, H. K. (1963). Additive effect of sub-threshold concentrations of some organic compounds associated with food aromas. *Nature*, **200**, 1288-1289.
- HAHN, H., & ULBRICH, L. (1948). Eine systematische Untersuchung der Geschmacksschwellen. *Pflügers Archiv*, **250**, 357-384.
- KROEZE, J. H. A. (1989). Is taste mixture suppression a peripheral or central event? In D. G. Laing, W. S. Cain, R. L. McBride, & B. W. Ache (Eds.), *Perception of complex smells and tastes* (pp. 225-243). San Diego, CA: Academic Press.
- KROEZE, J. H. A., & BARTOSHUK, L. M. (1985). Bitterness suppression as measured by split-tongue taste stimulation in humans. *Physiology & Behavior*, **35**, 779-783.
- LASKA, M., & HUDSON, R. (1991). A comparison of the detection thresholds of odour mixtures and their components. *Chemical Senses*, **16**, 651-662.
- LUCE, R. D. (1993). *Sound and hearing: A conceptual introduction*. Hillsdale, NJ: Erlbaum.
- MCBRIDE, R. L. (1989). Three models for taste mixtures. In D. G. Laing, W. S. Cain, R. L. McBride, & B. W. Ache (Eds.), *Perception of complex smells and tastes* (pp. 265-282). San Diego, CA: Academic Press.

- MCBRIDE, R. L., & JOHNSON, R. L. (1987). Perception of sugar-acid mixtures in lemon juice drink. *International Journal of Food Science & Technology*, **22**, 339-408.
- MOSKOWITZ, H. R. (1972). Perceptual changes in taste mixtures. *Perception & Psychophysics*, **11**, 257-262.
- PANGBORN, R. M., & TRABUE, I. M. (1967). Detection and apparent taste intensity of salt-acid mixtures in two media. *Perception & Psychophysics*, **2**, 503-509.
- PATTERSON, M. Q., STEVENS, J. C., CAIN, W. S., & COMETTO-MUÑIZ, J. E. (1993). Detection thresholds for an olfactory mixture and its three constituent components. *Chemical Senses*, **18**, 723-734.
- PFÄFFMANN, C., BARTOSHUK, L. M., & MCBURNEY, D. H. (1971). Taste psychophysics. In L. M. Beidler (Ed.), *Handbook of sensory physiology* (pp. 75-101). New York: Springer-Verlag.
- SCHIFFERSTEIN, H. N. J., & FRIJTERS, J. E. R. (1992). Two-stimulus versus one-stimulus procedure in the framework of functional measurement: A comparative investigation using quinine HCl/NaCl mixtures. *Chemical Senses*, **17**, 127-150.
- SCHIFFERSTEIN, H. N. J., & FRIJTERS, J. E. R. (1993). Perceptual integration in heterogeneous taste percepts. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 661-675.
- SMITH, D. V., & MCBURNEY, D. H. (1969). Gustatory cross-adaptation: Does a single mechanism code the salty taste? *Journal of Experimental Psychology*, **80**, 101-105.
- STEVENS, J. C., & CAIN, W. S. (1993). Changes in taste and flavor in aging. *Critical Reviews in Food Science & Nutrition*, **33**, 27-37.
- STEVENS, J. C., CAIN, W. S., DEMARQUE, A., & RUTHRUFF, A. M. (1991). On the discrimination of missing ingredients: Aging and salt flavor. *Appetite*, **16**, 129-140.
- STEVENS, S. S. (1957). On the psychophysical law. *Psychological Review*, **64**, 153-181.
- WETHERILL, G. B., & LEVITT, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical & Statistical Psychology*, **18**, 1-10.
- ZEKI, S. (1993). *A vision of the brain*. Oxford, U.K.: Blackwell Scientific Publications.

NOTE

1. It is near universal practice in threshold measurement to operate statistically on the logarithms of stimulus values. For visual quantities, this is usually explicitly stated (e.g., "log millilamberts"); for hearing, one uses the decibel notation, which is, of course, logarithmic; for the chemical senses, one typically specifies dilution steps. Since a dilution step is defined as a ratio, operating on dilution steps is like operating on logarithms of quantities or on decibels. This implies that statistical operations on dilution steps, such as means or standard deviations, are essentially geometric operations. In the instances in this paper where operations are done on stimulus values specified in terms of molar concentration, rather than dilution step, these are performed geometrically. That is, one takes the logs of the molarity first, then performs the statistical operation (e.g., a mean, standard error), and then takes the antilogarithm. The result is that all the statistical operations in this study are uniform, geometric, and in accordance with conventional practice.

(Manuscript received March 15, 1994;
revision accepted for publication July 12, 1994.)