Sensory Evaluation of Character Impact Components in an Apple Model Mixture

Johannes H.F. Bult[1,4], Hendrik N.J. Schifferstein[2], Jacques P. Roozen[1,4], Estanislau Dalmau Boronat[1], Alphons G.J. Voragen[1] and Jan H.A. Kroeze[3,4]

[1]Department of Agrotechnology and Food Sciences, Laboratory of Food Chemistry, Wageningen University and Research Centre, Wageningen, [2]Department of Industrial Design, Delft University of Technology, Delft, [3]Psychological Laboratory, Utrecht University, Utrecht and [4]Wageningen Taste and Smell Centre, Wageningen, The Netherlands

Correspondence to be sent to: A.G.J. Voragen, WUR, Department of Agrotechnology and Food Sciences, Laboratory of Food Chemistry, Biotechnion, PO Box 8129, 6700 EV, Wageningen, The Netherlands. e-mail: fons.voragen@chem.fdsci.wag-ur.nl

Abstract

Food aromas generally are complex mixtures of volatiles. In the present study, we investigated the joint effects of hexyl acetate, *trans*-2-hexenal and 1-hexanol on the multi-attribute perception of an apple aroma. The first two substances were identified earlier as positive contributors to the apple aroma (high character impact), whereas the third component was identified as an irrelevant or negative contributor (low character impact). Aroma quality was quantified using a set of eight graphic rating scales. All three components had significant effects on the aroma profiles. These effects consist mainly of an effect of each component on the attribute that described its individual character and an effect of all three components on ratings on the main character attribute 'apple'. As expected, the high impact components increased 'apple' ratings, whereas the low character impact component decreased 'apple' ratings. Furthermore, intensity ratings on the attribute that corresponded with the odour of the low impact component were suppressed by the presence of high impact components. These results indicate that the contributions of odorants to the mixture's aroma are not linear combinations of separate odour intensities, because sensory interactions were observed. In addition, humans detect components in complex mixtures more accurately than studies on identification performance have suggested. We conclude that for an adequate assessment of the effects of multiple mixture components on changes in aroma perception, it is sufficient to employ multiple response scales measuring intensities of attributes that are distinctive with respect to the expected qualitative changes. Results of this approach should be subjected to multivariate methods of statistical analysis.

Introduction

Food aromas generally comprise extensive mixtures of volatile constituents. A large number of these constituents produce odours if presented alone at similar concentration levels. Others, however, might not produce noticeable odours at all. An important objective in aroma research is to minimize the number of components in a modelled aroma by selecting those volatile components that contribute most to the original aroma. In general, this selection is made on the basis of two sensory properties: the relative perceived intensity of that component presented in isolation and the extent to which its character resembles the quality of the particular food aroma. Both higher perceived intensity and higher typicality of a component's odour quality result in a higher character impact of the component on the mixture's aroma (Buttery and Ling, 1998).

Character impact components (CICs) are usually identified by sensory analysis of mixture constituents after decomposing the mixture by gas chromatography (Dravnieks and O'Donnell, 1971). The constituents are evaluated with regard to their unmixed odour qualities and intensities. In doing so, one disregards that the contribution of an odorant to the mixture's aroma depends not only on its sensory characteristics when presented in isolation, but also on sensory interactions that occur when the odorant is perceived in the presence of other components.

Perceptual mixture interactions

In olfaction, partial mutual masking of mixture components is the most commonly observed interaction, even in mixtures consisting of as few as two components (Cain, 1975; Laing and Willcox, 1987; Lawless, 1987; Laing *et al.*, 1994). If one knows a component's psychophysical function, which relates component concentration to its perceived intensity, the intensity of a 'mixture' of that component with itself can be predicted from the sum of the two respective concentrations. However, the intensities of binary

mixtures of two different supra-threshold components are often lower than expected on the basis of their respective psychophysical functions (Moskowitz and Barbe, 1977; Berglund and Olsson, 1993a,b). On the other hand, indications of synergetic effects were observed when mixing sub-threshold components (Guadagni *et al.*, 1963; Laska and Hudson, 1991).

The mutual perceptual suppression of odorant intensities in multi-component mixtures is often observed (Moskowitz and Barbe, 1977; Moskowitz, 1979). Cain hypothesized that the masking power of supra-threshold odorants is positively related to either the chemical or the perceptual complexity of the masker (Cain, 1975). In the taste modality, the masking power of two substances in concert was indeed observed to be larger than the masking power of each of the substances alone (Stevens and Traverzo, 1997). In the case of olfaction, Laing and co-workers demonstrated that humans perform increasingly worse with increasing numbers of masking components when they are asked to identify odorous constituents in mixtures (Laing and Francis, 1989; Laing and Glemarec, 1992; Livermore and Laing, 1996; Jinks and Laing, 1999). Even the seemingly easy task of identifying the qualities of odorous constituents in binary mixtures yields probabilities of correct detections far below perfection (Olsson, 1994). A similar relationship between mixture complexity and masking power was observed when the mixture components themselves were multicomponent mixtures, each mixture representing a familiar object odour (Livermore and Laing, 1998).

Although humans experience great difficulties in recognizing the contribution of single components to the aroma of complex mixtures, they are able to discriminate between complex mixtures of odorants that are identical except for one component (Laska and Hudson, 1992). This can be explained by assuming that some or all of the odorants in a complex mixture blend perceptually into an intrinsically new aroma (Livermore and Laing, 1998). The omission of components from a complex mixture may then be detectable as a change in aroma quality, but not as an omission as such. Food aromas as well as many other object-related aromas generally consist of complex mixtures of odorants that, nonetheless, are perceived as homogeneous aroma blends. It is, therefore, rather speculative to assume that omitting components from a mixture would only affect the perceived intensity of their respective characters in the aroma of the mixture.

In recent years, scholars at the Deutsche Forschungsanstalt für Lebensmittelchemie have recognized the relevance of studying the contribution of CICs in the mixture. In a number of studies they evaluated the impact of components on mixture aroma by assessing the effect of omitting these from the mixture (Blank *et al.*, 1992; Guth and Grosch, 1994; Schieberle and Hofmann, 1997; Reiners and Grosch, 1998). In a study on the aroma of french fries (Wagner and Grosch, 1998), the authors determined the components with

high ratios of mixture concentration versus detection threshold concentration, called odour activity values (OAVs). Omitting these supra-threshold components from a model mixture often resulted in a significant discrimination of the aromas of the reduced versus the complete mixture. When reduced and complete mixtures were significantly perceived as different, panellists characterized the aroma qualities of the mixtures by rating intensities of attributes describing the odours of mixture components. In this study, omitting the component with the second highest OAV from the mixture was not detected in the discrimination task. After subsequent omission of additional components, however, a number of panellists gave higher ratings on the attribute describing the component with the second highest OAV. This *post hoc* evaluation suggests that this component was perceived only after it had been released from suppression due to a number of masking components. Indications for 'release from suppression' effects were also found in similar studies on an Arabica coffee model (Czerny *et al.*, 1999) and a white wine model (Grosch, 2000).

Evaluating aroma differences

The effect of component concentration on aroma quality can be quantified by measures of discriminability between mixtures, by similarity ratings, or by ratings on attributes describing the aromas. However, these methods vary with respect to their sensitivity to differences in aroma quality and their ability to characterize aroma quality.

If one only wishes to test hypotheses with respect to perceptual discriminability of stimuli, discrimination tasks may suffice. Trials in discrimination tasks yield binary-scaled results: a subject either does or does not distinguish correctly between differing stimuli. Proportions of correct stimulus discriminations can be calculated from repeated stimulus comparisons and are tested against the expected chance proportion of a correct discrimination. A psychophysical application of this method conceives the probability of correctly detecting a difference between stimuli as a measure of the sensory difference between stimuli (Thurstone, 1927; Frijters, 1980). It expresses sensory difference as the perceptual distance between stimuli on an arbitrary sensory continuum. Such measures of perceptual distance have been used to detect mixture interactions (Lawless and Schlegel, 1984).

Similarity ratings are used to measure the degree of sensory similarity between stimuli using discrete or continuous rating scales. Hence, the rated dissimilarity of stimuli can be conceived of as a measure of the distance between stimulus representations on an arbitrary sensory continuum. This sensory continuum may either represent a quality continuum or an intensity continuum, depending on the nature of the difference between stimuli.

Although discrimination tasks and similarity ratings are sensitive in detecting differences in both intensity and quality, they do not allow for semantic interpretations of

results in terms of odour quality characterizations. Therefore, to study the qualitative nature of sensory interactions, methods that directly address odour quality are needed. Attribute ratings reflect perceived intensities of odour characteristics indicated by odour quality descriptors. This makes attribute ratings an adequate tool to study and describe mixture interactions in both qualitative and quantitative terms. However, when the descriptor set is not distinctive with respect to the characteristic on which the stimuli differ, attribute ratings may have less discriminative power than similarity ratings (Callegari *et al.*, 1997). This may explain why Lawless and Schlegel (Lawless and Schlegel, 1984) found a taste–odour interaction in mixtures with variable sucrose and citral concentrations when using sensory distances calculated from discrimination task results, whereas no interaction was observed when attribute ratings were used. Intensity ratings on the attributes 'sweetness' and 'lemon odour' were merely statistically additive for the used stimulus set. In a meta-analysis, Callegari *et al.* showed that 25–30 distinctive descriptors are needed to cover the perceptual space for olfaction alone. Therefore, in the cross-modal study of Lawless and Schlegel, a set of two descriptors may have lacked the discriminative power needed to measure interactions. Dravnieks and colleagues (Dravnieks *et al.*, 1978) showed that over different panels, similarity ratings were at least as consistent as measures derived from attribute ratings. Summarizing, discrimination tasks and similarity ratings may be more sensitive or reliable methods to detect mixture interactions than attribute ratings, especially if selected descriptors are not distinctive. Nevertheless, the latter are to be preferred if the qualitative nature of these interactions should be assessed, provided that these are distinctive with respect to the characteristics on which stimuli differ.

Where no perceptual blending occurs in a mixture of odorants, we expect that the effect of changing a constituent's concentration in that mixture is best reflected by ratings on the component's corresponding descriptor. Panellists can generate these descriptors on presentation of the unmixed odorants, in which case these odorants can be used as standards to train panellists on the use of descriptors. This helps to align panellists' conceptual representations of attributes and, hence, may improve consistency of panel responses (O'Mahony, 1991; Lesschaeve and Sulmont, 1996). A descriptor set so designed may include as many descriptors as there are components in the mixture, which will often be less than the number of 25–30 recommended by Callegari *et al.* (Callegari *et al.*, 1997). Although a small selection of attributes covers merely a part of the olfactory universe, we argue that it is still a sensitive tool for describing the interactions in the mixture under investigation if component-derived descriptors are used.

Statistical interaction versus sensory interaction

When systematically manipulating the presence of a number of components in a mixture according to a factorial mixing design, one faces the task of deriving sensory mixture interactions from factorial plots of intensity measures. Note that this is not identical to identifying statistical interactions in the factorial plot. In general, statistical methods assume linear models relating dependent variables to independent variables. Psychophysical studies, however, have shown that the relationship between stimulus concentration and its perceived intensity rarely approaches linearity, but generally yields negatively accelerating curves fitting power functions with exponents ranging from 0.1 to 1.0 (Stevens, 1961; Cain, 1969; Baird *et al.*, 1996). It can be shown that, due to the non-linearity of psychophysical functions, factorial mixing plots of two-component mixture intensities will generally show converging lines, even when the 'constituents' are the same substance (De Graaf and Frijters, 1988). This may lead to a statistically significant interaction, while no sensory interaction is present. Only in those rare cases where the factorial mixing plot shows a set of diverging lines does a statistical interaction effect support a sensory interaction: a case of extremely strong synergism (De Graaf and Frijters, 1988; Schifferstein and Frijters, 1990; Schifferstein, 1995). In addition, sensory interactions are evident when a component suppresses an aspect of the mixture's aroma to which the component does not itself contribute. This will appear as a significant, negative statistical main effect and/or interaction effect of the suppressing component on intensity ratings. In the latter case, the suppressive effect does not necessarily coincide with a significant statistical interaction, although it does concern a sensory interaction. In this paper we will first report the outcomes of statistical tests and, subsequently, discuss these outcomes in terms of their implications with respect to sensory interactions.

The present study

In the present study, we investigated whether and how sensory interactions affect the perception of CICs in a complex mixture of odorants that observers recognize as a natural food aroma. To study the contribution of different CICs, we omit CICs systematically from the mixture. If a component's odour does not blend into an aroma at all, its impact can be measured using intensity ratings on its respective quality descriptor. Suppression of this component's intensity by other components can be measured accordingly. If, at the other extreme, all components contribute to one aroma blend, the main character descriptor of the aroma can be used to measure the impact of constituent components. The omission of a CIC should then reduce intensity ratings on the main character descriptor. In order to be able to describe aromas with different degrees of blending, we used a detailed aroma-profiling task involving both single component descriptors and a main character descriptor. If components contribute to the main aroma character and also remain individually distinguishable, effects on both

the main character descriptor and on specific quality descriptors will be observed.

Generally, food aromas are elicited by odorous components of varying odour intensities. Intensity is likely to be an important factor influencing the impact of an odorant on the mixture's aroma. In the present study, however, we wished to study processes involving odour quality only. To eliminate the effect of odour intensity, we matched the intensities of the three unmixed components under investigation before evaluating their effects in a multi-component model solution. These three components were selected according to their expected character impact: two components rated high on the target quality and one component rated low on this target quality.

Pilot study

Materials and methods

Subjects

Eighteen paid volunteers, five men and 13 women, served as subjects. They were recruited from the local Wageningen community and were selected on the basis of their ability to generate and use refined odour attributes. In addition, they showed high inter-subject coherency in the use of graphic rating scales. This implied that subjects generated inter-subject-correlating profiles when they rated various aroma intensities. All subjects were experienced olfactory panellists (Bult *et al.*, 2001), but they were naïve with respect to the objectives of the experiment. Their ages ranged from 19 to 51 (average 29 years). All were non-smokers and none had

any history of olfactory dysfunction. Subjects were in good health during the experiments. All gave written informed consent.

Stimuli

The aroma model that is used in this study was derived from a headspace sample of fresh apple juice earlier at this laboratory. Although the model consists of a limited number of components, it was recognized and described as apple by 13 out of 23 subjects upon presentation of the olfactory stimulus and without any extra information being given (Bult *et al.*, 2001). As identification performance for many common odours is ~50% (Cain, 1979), this model was deemed appropriate for the present study. Although we expect that more authentic apple aromas can be made, the aroma model in the present study may validly be assumed to represent a recognizable food aroma.

An apple reference stimulus was prepared from 10 components. Its composition was largely identical to that of the original apple mixture (Table 1). However, ethanol was excluded from the original mixture because it was of sub-threshold concentration, even after substantially increasing its concentration. Pre-testing revealed that the ethanol component did not induce any consistent olfactory sensations. Furthermore, the propyl propanoate concentration in the mixture was raised by a factor of five to enable a more accurate stimulus preparation, thus improving stimulus reliability. Since this component still had a low intensity in the given concentration, we assumed that this alteration did

Table 1 Substances used for the stimuli with their nominal purities and concentrations

Component	Nominal purity (%)	Attributes generated by the sensory panel (translated)	Concentration in water in pilot study (μl/l)	Base mixture component (base) or additional component (add.)	Concentration in water for apple reference aroma in pilot study (μl/l)	Concentration in water for profiling experiment (μl/l)
1-Butanol[c]	>99.5	sour, dairy	2000	base	15	30
2-Methyl-1-butyl acetate[a]	>94	sour hard-boiled candy[d]–glue	120	base	7.5	15
Butyl acetate[c]	>99	nail polish	400	base	12.5	25
Hexanal[c]	>98	macaroon[e]–hedge	25	base	12.5	25
Isobutyl acetate[a]	>99	sweet–lacquer	200	base	0.75	1.5
Propyl acetate[b]	>99	fruity–acetone	125	base	2.5	5
Propyl propanoate[c]	>99	spicy–eggnog	200	base	2.5	5
1-Hexanol[c]	>98	nuts–musty	300	add.	10	20/300[f]
Hexyl acetate[a]	>99	pear–apple	15	add.	7.5	15
Trans-2-hexenal[b]	>99	bittersweet–rum	80	add.	40	80

Obtained from [a]Aldrich, [b]Janssen Chimica, [c]Merck.

[d]Sour hard-boiled candy is a popular sweet in the Netherlands, where it is referred to as 'zuurtjes'.

[e]A macaroon is a cookie that has bitter almonds as its major flavour.

[f]1-Hexanol was present in both 20 μl/l and 300 μl/l concentrations (see text).

not have a significant impact on the character of the mixture aroma.

In addition to the mixture, 10 one-component stimuli were prepared from each of the 10 components in the apple model mixture. To obtain equi-intense stimuli, the concentration of each singular solution as well as the concentrations in the complete apple mixture were raised, so that all intensities matched the sensory intensity of an 80 µl/l solution of *trans*-2-hexenal. This was done in a preliminary study employing four faculty members. The resulting composition of the apple reference stimulus is given in the sixth column of Table 1. The concentrations of the singular solutions are given in the fourth column of Table 1. Mixtures and single-component dilutions were prepared using distilled water. All stimuli were prepared at least 2 h and not earlier than 26 h before presentation. Stimulus solutions were stored in the dark at 4°C and were presented at ambient temperature (21 ± 1°C).

Procedure

Stimuli were presented in 200 ml glass jars, closed by low-odour plastic screw caps, which could be opened by one simple twist. Each jar contained 10 ml of solution. To prevent volatile components from migrating from the screw caps to the headspace, these two phases were separated with aluminium foil. The subjects had to open each stimulus jar by unscrewing the cap, while keeping the jar just underneath their noses. Responses were to be given after taking a few short sniffs.

In the first session, the subjects generated odour attributes individually for all 10 unmixed stimuli. In the second part of this session, these attributes were discussed in a plenary meeting of all subjects. Consensus on the use of attributes was reached after plenary consultation in the second session.

At the start of the third session, the model mixture with the apple aroma was presented as the reference stimulus for 'apple' aroma. Subsequently, the 10 singular component solutions were presented. Of each solution, subjects rated the intensity of its 'apple' character on 150 mm scales printed on paper, labelled 'no apple' at the left end and 'very much apple' at the right end. The two components that scored highest and the one that scored lowest on 'apple' were selected as, respectively, two CICs and one non-character impact component (non-CIC). Note that the definition of character impact used here is based on quality rather than perceived intensity, since all stimuli were approximately equally intense.

Statistical analysis

Ratings for all 10 stimuli were collected within subjects. To compensate for idiosyncratic scale usage, attribute ratings were normalized to obtain equal means and standard deviations for each subject. For convenience, the complete data set for the group of subjects was transformed linearly in order to obtain a group score range from 0 to 100. This resulted in an average 'apple' value of 39.1 (SD = 22.4) for every subject, over 10 evaluated samples. Throughout the paper we used SPSS, version 7.5.2 (1997; SPSS, Chicago, IL) for data analyses and 0.05 as the level of significance.

Results

Averaged normalized 'apple' scores (±SE) are plotted for each component in Figure 1. Analysis of variance revealed a significant effect of 'component' on 'apple' responses [$F_{(9,179)} = 10.9$, $P < 0.01$]. Subsequent *post hoc* testing, using Duncan's multiple range statistic, showed that the 10 'apple' scores could be grouped in four clusters of not significantly different means (Figure 1). From the cluster of components rated highest on 'apple', hexyl acetate and *trans*-2-hexenal were selected as CICs, whereas 1-hexanol was selected from the lowest ranking cluster as a non-CIC.

Main Experiment

Materials and methods

Subjects

Eighteen paid volunteers, five men and 13 women, served as subjects for the main experiment. This group was identical to the group described in the pilot study, except for one female subject who was replaced by another female subject. The new subject met the criteria for admission to the panel as described for the pilot study. Ages ranged from 19 to 51 years and the average age was 30 years.

Stimuli

Ten different stimulus mixtures were prepared from the 10 selected components by systematically adding combinations of the two CICs and the non-CIC to a base solution of seven components in distilled water. Concentrations of the
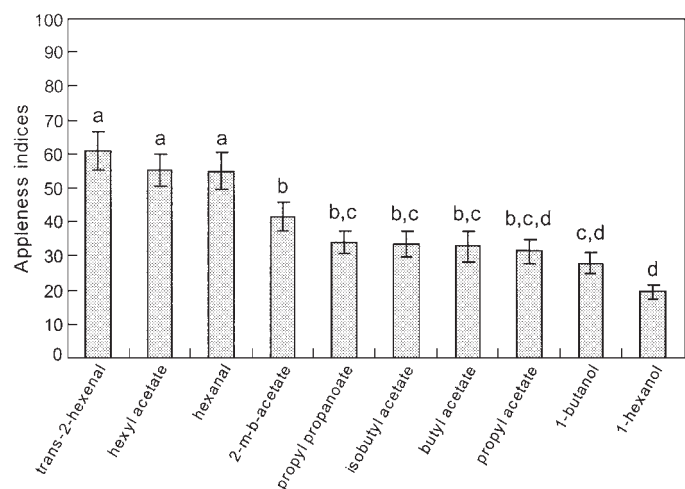

Figure 1 Normalized (0–100) apple ratings for 10 equi-intense components, expressed as mean (±SE) transforms of 150 mm scale ratings. Components are grouped in homogeneous groups according to the results of Duncan's multiple range test on 'apple' ratings. Identical letters indicate homogeneous groups.

base solution components and the additives are given in the last column of Table 1. The three additives—*trans*-2-hexenal, hexyl acetate and 1-hexanol—were added in singular, binary, or ternary combinations, thus resulting in 2 (presence of hexyl acetate = HYL) × 2 (presence of 1-hexanol = HOL) × 2 (presence of *trans*-2-hexenal = HAL) = 8 mixtures (Table 2). To match the unmixed intensities of hexyl acetate and *trans*-2-hexenal, the 1-hexanol component concentration had been increased to 300 ml/l relative to its original 20 ml/l concentration in the apple mixture. In addition, two mixtures containing the original—lower—concentrations of 1-hexanol were included in the stimulus set. One mixture consisted of the base mixture with only the low concentration 1-hexanol added (No. 9). The other also contained the two CICs hexyl acetate and *trans*-2-hexenal (No. 10, the original apple aroma).

Procedure

Subjects performed a descriptive analysis of the aromas of the 10 different stimuli. The attributes that were generated for the 10 unmixed substances were reduced to seven by letting each subject select three attributes that they considered the least appropriate descriptors for the full apple mixture. The attributes that were selected most frequently were discarded. In addition, an 'apple' attribute was included. Consequently, the eight attributes used in the descriptive analysis were: 'sour hard-boiled candy–glue'; 'macaroon–hedge'; 'sweet–lacquer'; 'fruity–acetone'; 'nuts–musty'; 'pear–apple'; 'bittersweet–rum' and 'apple'. The attribute names are translations of the Dutch terms used. Reference stimuli for the eight attributes (the apple mixture plus the respective components at concentrations identical to the pilot study) were presented prior to every experimental session in order to refresh odour-attribute associations. The use of the attributes 'apple' and 'pear–apple' may seem confounding because of their similarity. Subjects, however, perceived the respective qualities differently (see also Figures 2 and 3) and they considered these attributes the most appropriate for these aromas.

Stimulus preparation and presentation proceeded as described in the pilot study. One session lasted 40–50 min. Laboratory conditions conformed to the ISO 8589 standard (International Organization for Standardization, 1987). During a session, subjects were seated in separate booths. The uniform stimulus jars were coded with randomly generated three-digit codes and they were presented in random order, each individual receiving a separate order. Subjects were instructed to rate attribute intensities on eight linear 150 mm graphic rating scales that were presented on a laptop computer screen (Compaq Contura 80386 DX 25 MHz with monochrome display), using the left button of a two-button computer mouse. Between two stimuli, they waited for at least 1 min, which was computer-paced. After completing two training series in the first session, subjects completed nine experimental series of 10 stimuli each in three separate sessions. Consequently, nine evaluations of every distinct stimulus were recorded for every subject.

Statistical analysis

Data from the training sessions were discarded. Since no significant systematic changes in responses over sessions were observed, ratings were averaged over the nine repeated experimental sessions. Thus, data analyses were performed on averaged intensity scores on eight attribute variables for 10 different stimuli per subject.

Perceived aroma quality is reflected in the aroma profile. This does not imply that specific alterations of single component concentrations are reflected exclusively in attribute ratings of their respective accompanying attributes. Therefore, we initially tested for differences between complete profiles due to stimulus composition by doubly multivariate repeated measures analyses of variance—ANOVA (Stevens, 1996). These analyses permit simultaneous multivariate analyses of results on a set of dependent variables according to a repeated-measures design. The approach of initially performing a multivariate analysis also guards against spurious effects due to the increased overall significance level that results from performing successive univariate tests. Since the experiment had a fractional factorial design (10 categories from a 12 category full factorial design), the analysis was split into two consecutive multivariate analyses. First, CIC and non-CIC effects were tested in a 2 × 2 × 2, HOL (not present versus high concentration) × HYL × HAL design. Subsequently, the influence of all three

Table 2   Composition of stimulus mixtures derived from an apple model mixture

Constituents	Mixture composition (× = present)									
	1	2	3	4	5	6	7	8	9	10
Base mixture	×	×	×	×	×	×	×	×	×	×
Trans-2-hexenal (80 p.p.m.)		×			×	×		×		×
Hexyl acetate (15 p.p.m.)			×		×		×	×		×
1-Hexanol (300 p.p.m.)				×		×	×	×		
1-Hexanol (20 p.p.m.)									×	×

Mixture No. 1 constitutes the fully 'stripped' apple aroma, while mixture No. 10 represents the original apple aroma.
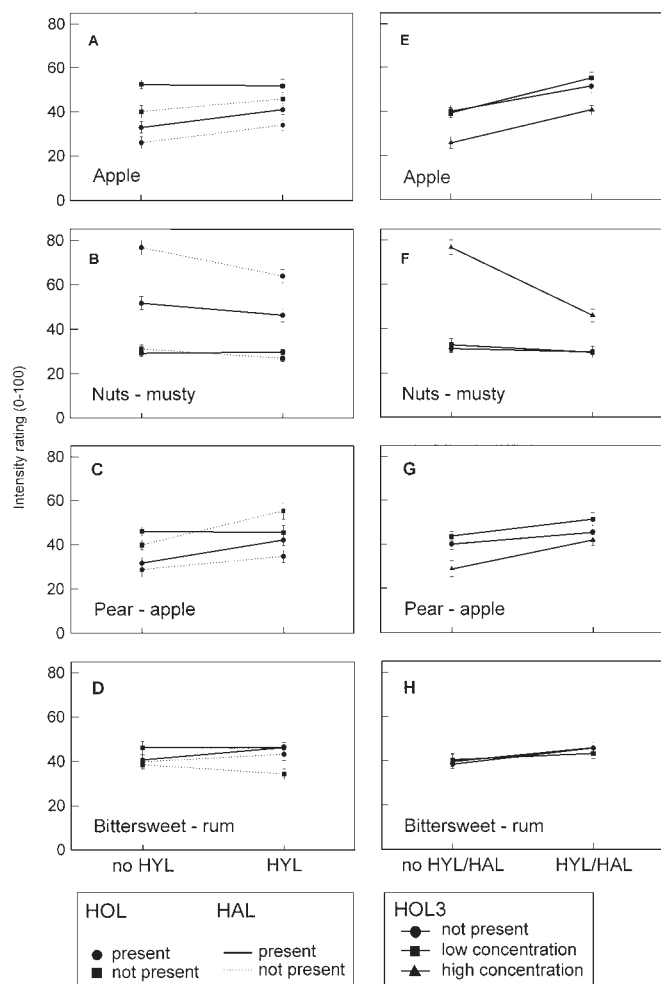
Figure 2   Mean intensity ratings (±SE) for the attributes 'apple', 'nuts–musty', 'pear–apple' and 'bittersweet–rum'. Panels **(A–D)** show the ratings for the base mixture with 1-hexanol—HOL (not present; present in high concentration), *trans*-2-hexenal—HAL (not present; present) and hexyl acetate—HYL (not present; present). Panels **(E–H)** show the ratings for the base mixture with 1-hexanol (not present; low concentration; high concentration) and a combination of *trans*-2-hexenal and hexyl acetate (not present; present).

1-hexanol concentrations (not present, low concentration, high concentration) and the presence or the absence of both *trans*-2-hexenal and hexyl acetate was tested in a 3 (HOL) × 2 (HYLHAL) design.

Any effects in the multivariate space indicate that aroma changes are perceived due to the addition or the omission of the CICs and the non-CIC. However, they do not give insight into the qualitative nature of the differences perceived. Because results on single attribute ratings may elucidate this in part, we proceeded with an analysis of single attribute data by univariate repeated measures ANOVA, as a *post hoc* test after significant multivariate effects are found. Multivariate *F*-values were calculated according to Pillai's trace criterion.

Results

The multivariate effects of CICs and the non-CIC on the eight attribute ratings in the HOL × HYL × HAL analysis were found significant for HOL [$F(8,10) = 4.80$, $P = 0.012$] and HAL [$F(8,10) = 13.12$, $P < 0.001$], whereas HYL [$F(8,10) = 2.48$, $P = 0.090$] failed to reach significance. A significant multivariate HOL × HAL interaction was also observed [$F(8,10) = 4.73$, $P = 0.013$]. No three-way interaction was found. The HOL3 × HYLHAL analysis of all three 1-hexanol levels yielded significant multivariate effects for HOL3 [$F(16,56) = 2.07$, $P = 0.023$] and HYLHAL [$F(8,10) = 3.97$, $P = 0.023$]. The HOL3 × HYLHAL interaction was also significant [$F(16,56) = 2.16$, $P = 0.017$].

Univariate repeated measures ANOVAs were performed for the main and the two-way interaction effects that were significant in the multivariate analyses. Effects were found on four out of eight dependent variables. All effects of component presence on aroma that were found in the multivariate analysis had counterparts in one or several of these univariate effects. These univariate effects, therefore, appear to explain the multivariate effects. Hence, further discussion of results will be restricted to the four dependent variables that showed significant effects.

Table 3 shows the ANOVA results for the HOL × HYL × HAL design grouped for each separate dependent variable and specified for separate sources of variance. Apple ratings, that reflect character impact of the three components in the mixture, show significant main effects of HOL, HAL and HYL. No interactions were found with respect to 'apple' ratings. In Figure 2A, the effects on aroma quality are illustrated. As may be expected on the basis of the nature of character impact components, addition of the CICs (HYL, HAL) to the base mixture increased 'apple' ratings. The addition of the non-CIC (HOL) decreased 'apple' ratings.

Most pronounced were the effects on the 'nuts–musty' ratings, which were significantly affected by all three components. The 'nuts–musty' attribute describes the character of the 1-hexanol component, which is reflected by significant higher 'nuts–musty' ratings at high HOL levels (Table 3, Figure 2B). Furthermore, a significant HOL × HAL interaction is found for 'nuts–musty' ratings. This interaction appears to be responsible for the multivariate HOL × HAL interaction found, since it is the only univariate interaction effect. It can be attributed to a masking influence of HAL on the 'nuts–musty' character introduced by HOL. The presence of *trans*-2-hexenal does not affect 'nuts–musty' ratings when 1-hexanol is not present. If 1-hexanol is present, however, the addition of *trans*-2-hexenal to the mixture suppresses the 'nuts-musty' character drastically (Figure 2B). Likewise, HYL, the other CIC, appears to exhibit a masking effect on the 'nuts–musty' character. Although the HOL × HYL interaction was not statistically significant, HYL had a significant main effect (Table 3) on 'nuts–musty'

Table 3   Repeated measures ANOVA of HOL × HYL × HAL

Attribute	Source	d.f.	F	P
Apple	HOL	1, 17	18.67	<0.001**
Apple	HYL	1, 17	7.16	0.016*
Apple	HAL	1, 17	16.92	0.001**
Apple	HOL × HAL	1, 17	0.03	0.865
Nuts–musty	HOL	1, 17	34.21	<0.001**
Nuts–musty	HYL	1, 17	15.81	0.001**
Nuts–musty	HAL	1, 17	21.26	<0.001**
Nuts–musty	HOL × HAL	1, 17	28.62	<0.001**
Pear–apple	HOL	1, 17	9.38	0.007**
Pear–apple	HYL	1, 17	10.61	0.005**
Pear–apple	HAL	1, 17	1.05	0.319
Pear–apple	HOL × HAL	1, 17	3.07	0.098
Bittersweet–rum	HOL	1, 17	0.04	0.836
Bittersweet–rum	HYL	1, 17	2.52	0.131
Bittersweet–rum	HAL	1, 17	4.59	0.047*
Bittersweet–rum	HOL × HAL	1, 17	1.94	0.182

Only the univariate results of the four attributes with significant effects are shown. *$P < 0.05$; **$P < 0.01$.

Table 4   Repeated measures ANOVA of HOL3 × HYLHAL for the same attributes as in Table 3

Attribute	Source	d.f.	F	P
Apple	HOL3	2, 34	12.13	<0.001**
Apple	HYLHAL	1, 17	19.22	<0.001**
Apple	HOL3 × HYLHAL	1, 17	0.70	0.499
Nuts–musty	HOL3	2, 34	32.34	<0.001**
Nuts–musty	HYLHAL	1, 17	23.86	<0.001**
Nuts–musty	HOL3 × HYLHAL	1, 17	17.17	<0.001**
Pear–apple	HOL3	2, 34	6.05	0.006**
Pear–apple	HYLHAL	1, 17	8.54	0.010**
Pear–apple	HOL3 × HYLHAL	1, 17	1.52	0.234
Bittersweet–rum	HOL3	1, 17	0.10	0.909
Bittersweet–rum	HYLHAL	1, 17	4.12	0.058
Bittersweet–rum	HOL3 × HYLHAL	1, 17	1.60	0.217

**$P < 0.01$.

HOL3 were due to the influence of the highest 1-hexanol concentration level.

ratings comprising a decrease in 'nuts–musty' ratings due to HYL (Figure 2B).

HYL had a significant effect on its character descriptor 'pear–apple'. As may be expected, this effect comprised an increase of 'pear–apple' ratings after adding hexyl acetate to the mixture. Furthermore, 'pear–apple' ratings decreased significantly when 1-hexanol was added to the mixture (Figure 2C).

Ratings on 'bittersweet–rum' also increased significantly when its characteristic component, *trans*-2-hexenal, was added to the mixture (Table 3, Figure 2D). No other effects were found for this descriptor.

Table 4 shows the ANOVA results for the HOL3 × HYLHAL 3 × 2 design, grouped for each separate dependent variable and specified for separate sources of variance. The results are similar to those presented in Table 3. High HOL levels suppress 'apple' and 'pear–apple' ratings and increase 'nuts–musty' ratings (Figure 2E,G,F, respectively). Ratings on 'apple' and 'pear–apple' increased when CICs (HYLHAL) were added (Figure 2E,G, respectively). The contribution of HOL3 to 'nuts–musty' is suppressed by HYLHAL (Figure 2F).

Applying orthogonal simple contrasts on HOL3 levels (none, low, high), comparing the levels 'none' to 'low', respectively 'none' to 'high', revealed significant effects for 'high' versus 'none' on ratings for 'apple' [$F(1,17) = 14.48$, $P = 0.001$], 'pear–apple' [$F(1,17) = 4.84$, $P = 0.042$] and 'nuts–musty' [$F(1,17) = 36.16$, $P < 0.001$]. No significant effects of 'low' versus 'none' were observed. Likewise, the HOL3 × HYLHAL interaction could be attributed to the interaction between 'none' versus 'high' (HOL) and HYLHAL. Therefore, all main and interaction effects of

Discussion

If only 'apple' ratings are taken into account, the CICs and non-CIC investigated here show statistical additivity. Adding HYL or HAL to the apple base mixture increases apple ratings. Whether HYL and HAL produce sensory hypo- or hyper-additivity cannot be concluded on the basis of the present data, since this requires more information on the form of the psychophysical functions of these two substances (De Graaf and Frijters, 1988; Schifferstein, 1995). The presence of 1-hexanol suppresses 'apple' ratings for all mixtures, reflecting a marked sensory interaction. Correspondingly, 1-hexanol suppresses mixture ratings on the 'pear–apple' attribute.

When the multivariate HOL × HAL interaction was investigated further by studying single attribute effects, this interaction could be attributed to *trans*-2-hexenal suppressing the 'nuts–musty' intensity at high 1-hexanol levels. In other words, the CIC *trans*-2-hexenal suppresses the contribution of 1-hexanol to the ratings on its corresponding attribute. This adds to the status of *trans*-2-hexenal as 'character impact component' since it suppresses part of the effect of a non-character impact component on the mixture's aroma. Summarized, the observed interactions show that sensory interactions among mixture components are present and that these interactions pertain to ratings on a number of attributes.

In experimental investigations of mixture aroma quality, a single attribute describing the main character of the aroma cannot sufficiently reflect contributions of all components to the aroma. Had, for instance, only 'apple' ratings been used in this study, then the important *trans*-2-hexenal × 1-hexanol interaction would have gone unnoticed. There-

fore, we argue that the use of multiple attribute ratings should be preferred to one-dimensional measures in food aroma studies. However, a limited set of eight attributes is rather small according to recommended 25–30 (Callegari *et al.*, 1997). Therefore, some concern is justified with respect to the validity of the operationalization of aroma quality in the present study. When using the common technique of odour profiling as an operationalization of aroma quality, one assumes a linear additive model for contributions of each attribute to the overall aroma. Therefore, an observed mutual suppression of components can not result from the chosen operationalization technique. Hence, we attribute the sensitivity for mixture interactions of our characterization method to the use of descriptors generated on the basis of the constituent odours. This enabled the training of attribute usage and allowed for direct measurement of mutual suppression of odorants in the mixture. Although the operationalization used here may not allow full representation of perceived aromas, it proved sufficient for the assessment of mixture interactions.

To eliminate the effect of odour intensity, the intensities of the two CICs and the non-CIC were matched before their contribution to a multi-component model solution was investigated. This involved raising the concentration of the non-CIC component. Although this may have altered the quality of the aroma, it was necessary to do so in order to be able to attribute effects exclusively to the influence of odour quality. Interestingly, the highest rating on the apple attribute was given for the original apple model (Figure 2E), in which the non-CIC was present in low concentration.

The two CICs and the non-CIC showed main effects on their corresponding odour attributes. From this, it can be concluded that these components did indeed influence the perceived aromas. More specifically, when the three studied components were added to the mixture, ratings for the three respective character descriptors increased significantly. This suggests that panellists were able to recognize the unique contribution of each of the three manipulated components to the mixture's aroma. This is surprising given that Laing and colleagues showed that humans have great difficulty recognizing as few as three or four components in mixtures containing up to five or six odorants (Laing and Francis, 1989; Laing and Glemarec, 1992; Livermore and Laing, 1996). This discrepancy in results is unlikely to be due to differing similarities between the odorants used in each study. Laing *et al.* employed dissimilar odorants, which should have maximized the number of correctly identified components, whereas at least one odorant in the present study was both structurally and perceptually similar to a CIC (i.e. hexanal and *trans*-2-hexenal). Also, differing mixture complexities cannot have caused this outcome since the mixture in the present study is more complex, creating a more difficult task.

An explanation for the seemingly enhanced performance of subjects in the present experiment can be found in the methodology employed. In the present study, the subjects were aided by being provided with specific descriptors that directed them in rating specific feature intensities. Subjects were not requested to focus on physical components, as was the case in the study by Laing *et al*. Furthermore, Laing *et al*. gave their subjects dichotomous decision options: an odorant is present in the mixture or it is not. Under this regime a subject has a complex task: he or she has to assess the intensity of component-specific contributions to the mixture and has to decide on the relevance of the perceived intensities to the question of whether components are present or not. In contrast, the present study employed continuous attribute scales that enabled subjects to express the intensity of sensations. No absolute decisions on presence had to be made.

Acknowledgements

References

Baird, J.C., Berglund, B. and **Olsson, M.J.** (1996) *Magnitude estimation of perceived odor intensity: empirical and theoretical properties*. J. Exp. Psychol. Hum. Percept. Perform., 22, 244–255.

Berglund, B. and **Olsson, M.J.** (1993a) *Odor-intensity interaction in binary mixtures*. J. Exp. Psychol. Human Percept. Perform., 19, 302–314.

Berglund, B. and **Olsson, M.J.** (1993b) *Odor-intensity interaction in binary and ternary mixtures*. Percept. Psychophys., 53, 475–482.

Blank, I., Sen, A. and **Grosch, W.** (1992) *Sensory study on the character-impact flavour compounds of dill herb (*Anethum graveolens *L.)*. Food Chem., 43, 337–343.

Bult, J.H.F., Schifferstein, H.N.J., Roozen, J.P., Voragen, A.G.J. and **Kroeze, J.H.A.** (2001) *The influence of olfactory concept on the probability of detecting sub- and peri-threshold components in a mixture of odorants*. Chem. Senses, 26, 459–469.

Buttery, R.G. and **Ling, L.C.** (1998) *Additional studies on flavor components of tortilla chips*. J. Agric. Food Chem., 46, 2764–2769.

Cain, W.S. (1969) *Odor intensity: differences in the exponent of the psychophysical function*. Percept. Psychophys., 6, 349–354.

Cain, W.S. (1975) *Odor intensity: mixtures and masking*. Chem. Senses Flav., 1, 339–352.

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

Callegari, P., Rouault, J. and **Laffort, P.** (1997) *Olfactory quality: from descriptor profiles to similarities*. Chem. Senses, 22, 1–8.

Czerny, M., Mayer, F. and **Grosch, W.** (1999) *Sensory study on the character impact odorants of roasted arabica coffee*. J. Agric. Food Chem., 47, 695–699.

De Graaf, C. and **Frijters, J.E.R.** (1988) *Assessment of the taste interaction between qualitatively similar-tasting substances: a comparison between comparison rules*. J. Exp. Psychol. Human Percept. Perform., 14, 526–538.

Dravnieks, A. and **O'Donnell, A.** (1971) *Principles and some techniques*

of high-resolution headspace analysis. J. Agric. Food Chem., 19, 1049–1056.

Dravnieks, A., Bock, F.C., Powers, J.J., Tibbetts, M. and **Ford, M.** (1978) *Comparison of odors directly and through profiling*. Chem. Senses Flav., 3, 191–225.

Frijters, J.E.R. (1980) *Psychophysical and psychometrical models for the triangular method*. Doctoral thesis, Utrecht University, Utrecht, The Netherlands.

Grosch, W. (2000) *Specificity of the human nose in perceiving food odorants*. In Schieberle, P. and Engel, K.H. (eds), The Proceedings of the 9th Weurman Flavour Research Symposium, Freising, Germany, 22–25 June 1999. Deutsche Forschungsanstalt für Lebensmittelchemie, Garching, pp. 213–219.

Guadagni, D.G., Buttery, R.G., Okano, S. and **Burr, H.K.** (1963) *Additive effect of sub-threshold concentrations of some organic compounds associated with food aromas*. Nature, 200, 1288–1289.

Guth, H. and **Grosch, W.** (1994) *Identification of the character impact odorants of stewed beef juice by instrumental analyses and sensory studies*. J. Agric. Food Chem., 42, 2862–2866.

International Organization for Standardization (ISO) (1987) Sensory Analysis—General Guidance for the Design of Test Rooms, ISO 8589, 1–9. ISO, Geneva.

Jinks, A. and **Laing, D.G.** (1999) *A limit in the processing of components in odour mixtures*. Perception, 28, 395–404.

Laing, D.G. and **Francis, G.W.** (1989) *The capacity of humans to identify odors in mixtures*. Physiol. Behav., 46, 809–814.

Laing, D.G. and **Glemarec, A.** (1992) *Selective attention and the perceptual analysis of odor mixtures*. Physiol. Behav., 52, 1047–1053.

Laing, D.G. and **Willcox, M.E.** (1987) *An investigation of the mechanisms of odor suppression using physical and dichorhinic mixtures*. Behav. Brain Res., 26, 79–87.

Laing, D.G., Eddy, A., Francis, G.W. and **Stephens, L.** (1994) *Evidence for the temporal processing of odor mixtures in humans*. Brain Res., 651, 317–328.

Laska, M. and **Hudson, R.** (1991) *A comparison of the detection thresholds of odour mixtures and their components*. Chem. Senses, 16, 651–662.

Laska, M. and **Hudson, R.** (1992) *Ability to discriminate between related odor mixtures*. Chem. Senses, 17, 403–415.

Lawless, H.T. (1987) *An olfactory analogy to release from mixture suppression in taste*. Bull. Psychonom. Soc., 25, 266–268.

Lawless, H.T. and **Schlegel, M.P.** (1984) *Direct and indirect scaling of sensory differences in simple taste and odor mixtures*. J. Food Sci., 49, 44–46, 51.

Lesschaeve, I. and **Sulmont, C.** (1996) *Defining odour descriptors with one or three standards*. 12th International Symposium on Olfaction and Taste, 25–31 August 1996, Zürich, p. 78.

Livermore, A. and **Laing, D.G.** (1996) *Influence of training and experience on the perception of multicomponent odor mixtures*. J. Exp. Psychol. Human Percept. Perform., 22, 267–277.

Livermore, A. and **Laing, D.G.** (1998) *The influence of chemical complexity on the perception of multicomponent odor mixtures*. Percept. Psychophys., 60, 650–661.

Moskowitz, H.R. (1979) *Utility of the vector model for higher-order mixtures: a correction*. Sensory Processes, 3, 366–369.

Moskowitz, H.R. and **Barbe, C.D.** (1977) *Profiling of odor components and their mixtures*. Sensory Processes, 1, 212–226.

O'Mahony, M. (1991) *Descriptive analysis and concept alignment*. In Lawless, H.T. and Klein, B.P. (eds), Sensory Science Theory and Applications in Foods. Marcel Dekker, New York, pp. 223–268.

Olsson, M.J. (1994) *An interaction model for odor quality and intensity*. Percept. Psychophys., 55, 363–372.

Reiners, J. and **Grosch, W.** (1998) *Odorants of virgin olive oils with different flavor profiles*. J. Agric. Food Chem., 46, 2754–2763.

Schieberle, P. and **Hofmann, T.** (1997) *Evaluation of the character impact odorants in fresh strawberry juice by quantitative measurements and sensory studies on model mixtures*. J. Agric. Food Chem., 45, 227–232.

Schifferstein, H.N.J. (1995) *Perception of taste mixtures.* In Doty, R.L. (ed.), Handbook of Olfaction and Gustation. Marcel Dekker, New York, pp. 689–713.

Schifferstein, H.N.J. and **Frijters, J.E.R.** (1990) *Sensory integration in citric acid/sucrose mixtures*. Chem. Senses, 15, 87–109.

Stevens, J. (1996) Applied Multivariate Statistics for the Social Sciences. Lawrence Erlbaum, Mahwah, NJ.

Stevens, J.C. and **Traverzo, A.** (1997) *Detection of a target taste in a complex masker*. Chem. Senses, 22, 529–534.

Stevens, S.S. (1961) *Toward a resolution of the Fechner–Thurstone legacy*. Psychometrika, 26, 35–47.

Thurstone, L.L. (1927) *A law of comparative judgement*. Psychol. Rev., 34, 273–286.

Wagner, R.K. and **Grosch, W.** (1998) *Key odorants of french fries*. J. Am. Oil Chem. Soc., 75, 1385–1392.