Chem. Senses 25: 429–443, 2000

REVIEW

Quantification of Odor Quality

Paul M. Wise, Mats J. Olsson[1] and William S. Cain

Chemosensory Perception Laboratory, Department of Surgery (Otolaryngology), University of California, San Diego, La Jolla, CA 92093-0957, USA and [1]Department of Psychology, Uppsala University, S-75142 Uppsala, Sweden

Correspondence to be sent to: William S. Cain, Chemosensory Perception Laboratory, Department of Surgery (Otolaryngology), Mail Code 0957, University of California, San Diego, La Jolla, CA 92093-0957, USA. e-mail: wcain@ucsd.edu

Abstract

The relationship between odor quality and molecular properties is arguably the most important issue in olfaction. Despite sophistication in the chemical characterization of molecules, accompanying perceptual characterization has had little quantitative usefulness, relying mostly on enumerative description. As a result of weak interest in the topic outside industry and little agreement regarding how to measure quality, the field of olfactory psychophysics has failed to develop a substantial database for odor quality and has offered little help to other researchers, e.g. neurobiologists, in choice of stimuli, interpretation of outcome or testable hypotheses. This review scrutinizes how psychophysicists and others have measured quality and offers criteria for useful techniques. Most measures have had a subjective component that makes them anachronistic with modern methodology in experimental behavioral science, indeterminate regarding the extent of individual differences, unusable with infrahumans and of unproved ability to discern small differences. Techniques based upon performance, rather than on the more common reporting of mental content, offer firmer possibilities for growth. These techniques inevitably tap the discriminative basis of perception. The nonsubjective techniques have high sensitivity, can have counterparts in infrahuman research, are suitable to examine individual differences and yield non-negotiable answers with potential archival value. Discriminative techniques have their limitations, too—principally excess sensitivity that abridges their use to comparisons between similar-smelling stimuli. Research has begun to extend that range and may overcome the limitation. Application of discriminative methods may have the side-effect of shifting focus in structure–activity research from searches for molecular least common denominators that underlie often vague similarity to the search for molecular properties of importance in discrimination of small differences.

Introduction

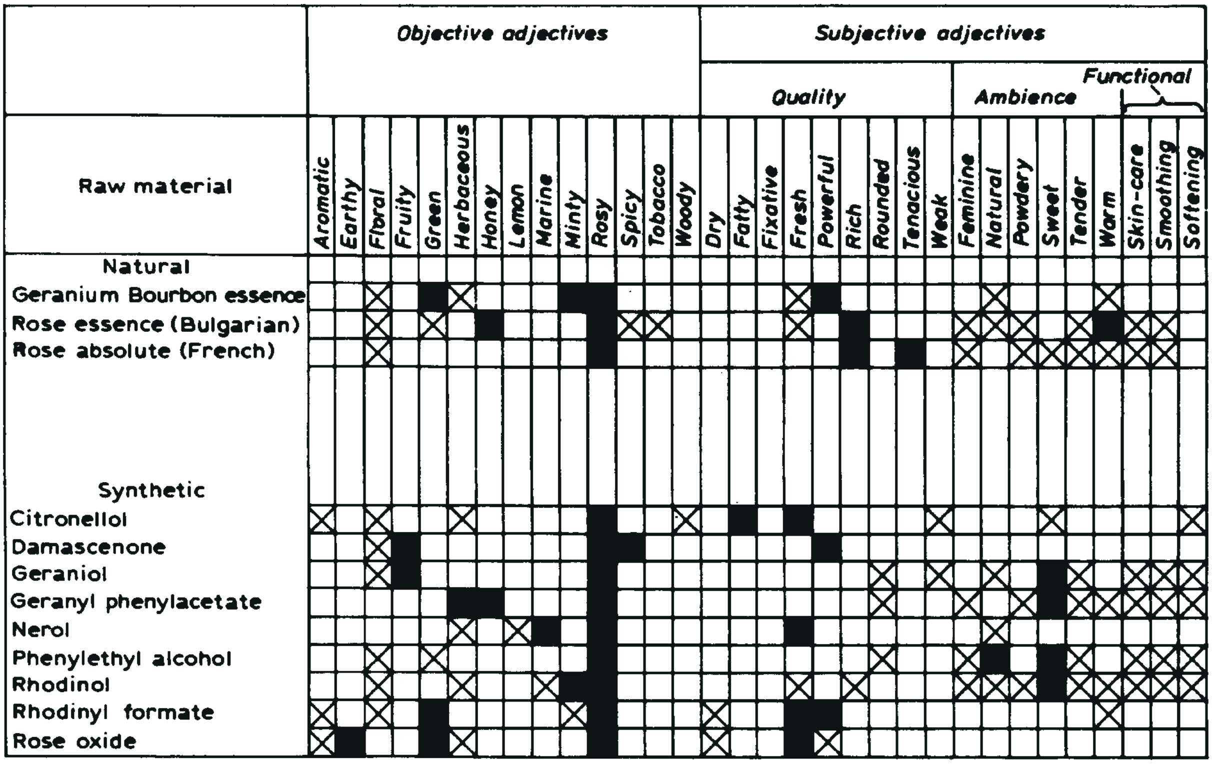
> Can you measure the difference between one kind of smell and another? It is very obvious that we have very many different kinds of smells, all the way from the odour of violets and roses up to asafetida. But until you can measure their likeness and differences you can have no science of odour. (Alexander Graham Bell, 1914)

In the many decades since Bell made his observation, no such science of odor has materialized. Scientists have neither measured likenesses and differences very effectively nor deciphered what causes them. Various notions concerning the relationship between properties of molecules and their corresponding odors have appeared, but none has attained acceptance as a legitimate theory (Cain, 1988; Rossiter, 1996, Chastrette, 1997). As Bell saw, it is axiomatic that any account of odor quality should develop around a corpus of measurements. Science makes incremental progress through cycles of data-collection followed by theorizing or model-building followed by more data-collection. In the case of odor, few trustworthy measurements of likenesses and differences have preceded theory and few have followed. Theory has accordingly had few data to explain and has stimulated little research.

To develop a corpus, the field needs trustworthy measuring techniques. This paper examines conditions for some choices.

Criteria for a measure

Techniques to measure quality could usefully meet three criteria: (i) they should resolve quite small differences since the clearest insights into molecular determinants of quality will undoubtedly come in the quest to account for small differences; (ii) they should produce indices with properties



	Objective adjectives														Subjective adjectives																	
															Quality									Ambience						Functional		
Raw material	Aromatic	Earthy	Floral	Fruity	Green	Herbaceous	Honey	Lemon	Marine	Minty	Rosy	Spicy	Tobacco	Woody	Dry	Fatty	Fixative	Fresh	Powerful	Rich	Rounded	Tenacious	Weak	Feminine	Natural	Powdery	Sweet	Tender	Warm	Skin-care	Smoothing	Softening
Natural																																
Geranium Bourbon essence			X		■	X				■	■							X	■						X				X			
Rose essence (Bulgarian)			X		X		■				■	X	X					X		■				X	X	X		X	■	X	X	
Rose absolute (French)			X								■									■		■		X		X	X	X	X	X	X	
Synthetic																																
Citronellol	X		X			X					■			X		■		■					X				X					X
Damascenone			X	■							■	■							■													
Geraniol			X	■							■										X		X		X		■	X		X	X	X
Geranyl phenylacetate						■	■				■										X			X		X	■	X	X	X	X	X
Nerol						X		X	■		■							■							X							
Phenylethyl alcohol			X		X						■										X			X	■		■	X		X	X	X
Rhodinol			X			X			X	■	■							X		X				X	X	X	■	X	X	X	X	X
Rhodinyl formate	X		X		■					X	■				X			■	■										X			
Rose oxide	X	■			■	X					■				X			■	X													

Figure 1 Illustration of odor characterization by a perfumer. Thiboud (Thiboud, 1991) used 18 'subjective' descriptors and 14 'objective' descriptors to characterize secondary notes (Xs) in rosy-smelling materials. He based his subjective description 'on the ambience created or brought on by these . . . products' (p. 255), whereas he based his objective description 'on the olfactory note of the raw material' (p. 255). The term 'subjective' follows normal usage in the text of the present piece, i.e. existing within the mind and incapable of external verification. By this standard, no descriptors qualify as objective. Reprinted from (Thiboud, 1991), table 11, p. 268, with kind permission from Kluwer Academic Publishers.

of interval-scale measurement (Stevens, 1950, 1960) to allow metric comparisons with molecular properties; and (iii) they should avoid subjectivity. The third criterion places the most serious constraint on the techniques, for almost all used through the years have a subjective aspect, a matter not always acknowledged explicitly (Figure 1).

Before moving on, definitions of objective and subjective (as the terms will be used in this piece) might prove useful. An objective assessment is based on performance, whereas a subjective assessment is based on a report of mental content. Suppose, for example, one wished to assess the qualitative similarity of two odors. One might determine how well a subject can tell the two odors apart (a performance-based, or objective, assessment) or ask the subject how different the odors seem (a report of mental content, or subjective, assessment).

In the world of applied sensory measurement, subjective characterization can serve satisfactorily to communicate the aromas of materials, foods, beverages, personal products, etc., and in this role deserve no particular disparagement (Swan and Burtles, 1980; Stone and Sidel, 1985). In order to characterize an aroma absolutely, rather than comparatively, some such description may seem unavoidable, but when used to test the determinants of quality, description may limit progress by providing outcomes sometimes as vague as Rorschach inkblots (Figure 2). Some techniques entail greater subjectivity than others and it would be imprudent to tar all with the same brush. Nevertheless, reliance on subjective techniques at all has at least three important liabilities:

1. Subjective answers are commonly circumstantial in that they depend on context, on experience, on attitude, on culture and on personal factors of unspecified origin. Technically, one may not dispute any person's description, for that would violate a private domain. Yet, by the same token, every description is open to contradiction and negotiation as just one person's opinion and hence not necessarily 'true'. When different persons characterize the same odor differently, the reasons are generally left unexplained. Descriptions often achieve stability only when obtained from groups of subjects. Even so, a group of Americans may agree on one description and a group of French people on another. Which group has yielded the objective answer?
2. Subjective measures can lead to rapid generation of data, but the speed comes with a price, namely an inability to give careful consideration to individual differences (even if such differences did reflect true differences in

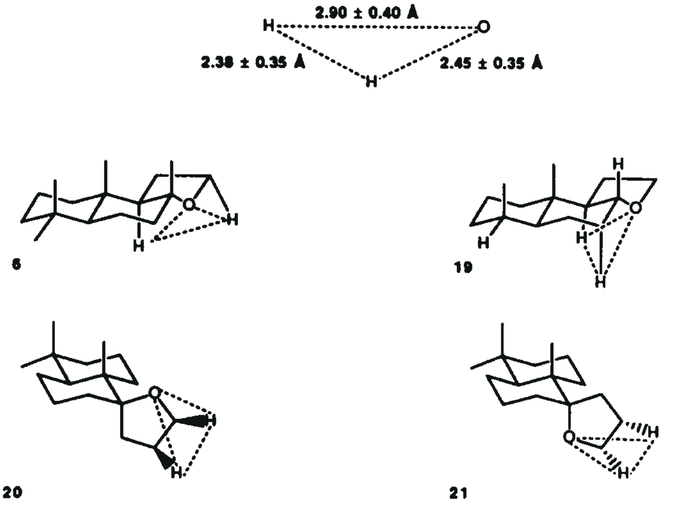


Figure 2 Most studies of structure–odor relationships aim at practical goals, i.e. to characterize interesting new fragrance and flavor materials. Consequently, studies focus on such stimuli as ambergris, sandalwood, musk, bell pepper and almonds. According to Ohloff and co-workers (Ohloff *et al.*, 1991), aged tincture of ambergris, a natural mixture, has the qualities: (i) wet, mossy forest soil; (ii) strong tobacco; (iii) balsamic, sandalwood-like; (iv) warm animal tonality of musk; (v) seaweed, ocean; and (vi) fecal. Scores of synthesized compounds have ambergris-like odors. Ohloff *et al.* noted that Ambrox (6) matched the first four tonalities, which made it very useful in perfumery. This characterization of Ambrox is an example of the value of enumerative description to professionals in the fragrance industry. In an effort to understand why substances structurally different from Ambrox also have ambergris-like odors, Vlad and colleagues proposed an 'ambergris triangle' of the dimensions shown above. The use of enumerative description so valuable for communication among professionals now invites the reductionistic conclusion that single molecules must behave like complex mixtures with different ingredients, now 'features'. Without any actual quantitative indication of how similar the compounds 6, 19, 20 and 21 smell to one another, to Ambrox (which had only four of the 'tonalities' of tincture of ambergris) or to ambergris itself, one cannot assess the validity of the triangle. This work, ostensibly on structure–activity relationships, exemplifies many studies in the literature where enumerative description has led to suggestions about the structural basis for one or another quality (Napolitano *et al.*, 1994, 1996). Such studies may occasionally provide useful information for fragrance professionals and may give some leads to the synthetic chemist, but have rarely yielded enduring scientific understanding. Reprinted from (Ohloff *et al.*, 1991), figure 3, p. 293, with permission kind permission from Kluwer Academic Publishers.

reception). As discussed below, subjective techniques, unlike nonsubjective techniques that measure performance, entail the collection of just a small amount of data per odor comparison. Once a subject describes an odor, such as ethyl butyrate, as fruity–fragrant–aromatic–sweet, that subject's job is typically done (Cain *et al.*, 1998). This holds even for, say, numerical ratings of similarity between odors. After one or two ratings for a particular pair of odors, a subject will remember the ratings and on subsequent trials merely repeat them, showing spuriously high repeatability with himself, though adding no new information.

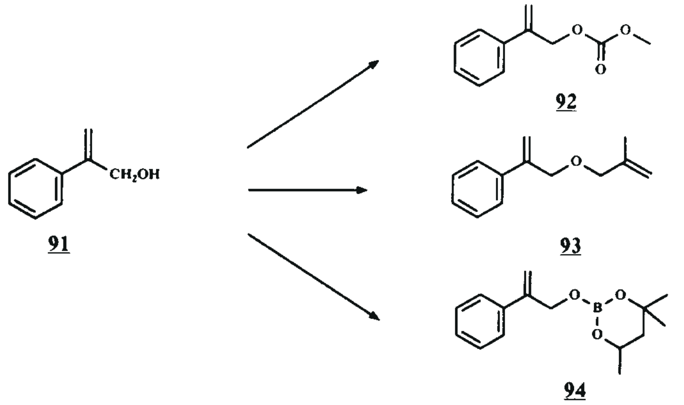


Figure 3 Derivatives of phenylallyl alcohol (91) smell rosy, but with other notes: compound 92 smells rosy with 'a lilac and spicy shadow', compound 93 smells rosy with a 'greener, ozone-like smell with fruity top note' and compound 94 smells like 'rose, cinnamon, carnation, spice, and lilac'. Such stimuli as these could intrigue a researcher who records neural potentials in, for example, the olfactory bulb. In this context, it might help to know that these compounds share a rosy characteristic. However, without quantitative data on perceived similarity, both for humans and for the model organism, neurophysiologists cannot test quantitative models of the coding of odor quality. Reprinted from (Fráter and Lamparsky, 1991), p. 587, with kind permission from Kluwer Academic Publishers.

3. Subjective techniques generally preclude isometric comparisons between the data of humans and that of other animals. Interspecies comparisons, combined with physiological data on the infrahuman, represent a historically important path to understanding. In light of the progress in the study of neural coding in olfaction (Mori and Shepherd, 1994; Yokoi *et al.*, 1995; Buck, 1996), lack of commensurate progress in the psychophysics of smell seems a badly missed opportunity. Psychophysical research could provide key information about odors in the neurobiological experiment (Figure 3).

How have odors been measured?

Classification

From the time of the ancient Greeks, scholars have grouped odors by qualitative resemblance (Beare, 1906; Zwaardemaker, 1925; Boring, 1942; Harper *et al.*, 1968; Cain, 1978). Such schemes can offer ordinal measurement. In the hierarchical categorization scheme that one can create as a dendrogram, similarity can be seen graphically to decrease ordinally the closer to the root one must progress to find a link between any two odors (Figure 4).

Categorization as a judgement may seem harmlessly nonsubjective, but a person's background and personal theory of reality can actually influence categories as much as, or even more than, basic sensory similarity (Smith and Medin, 1981; Murphy and Medin, 1985). Do people call various floral aromas by that name because they smell so much alike or because they seem to come from flowers? Do people call spicy aromas by that name because they smell

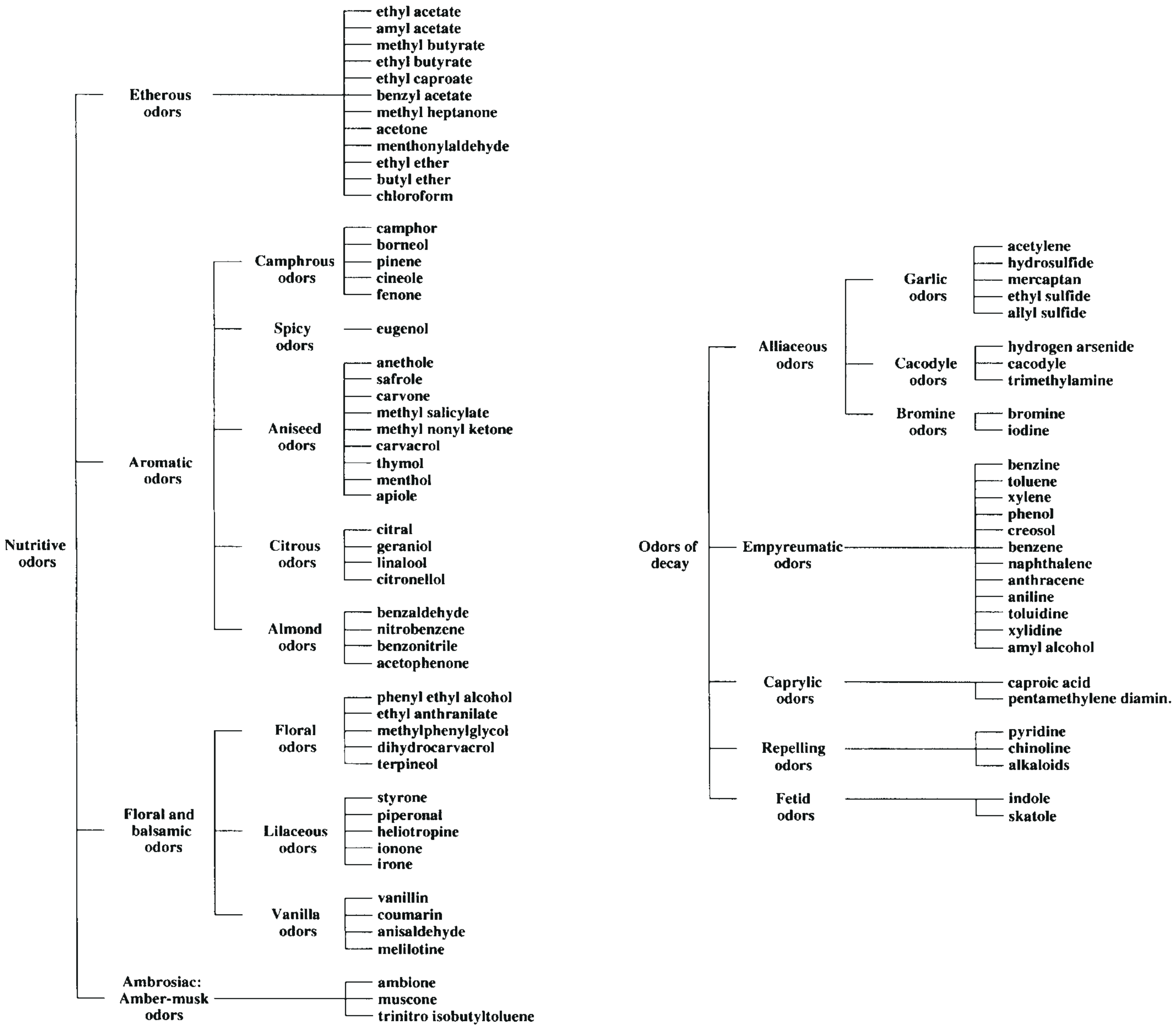


Figure 4 A dendrogram such as that constructed for Zwaardemaker's classification of odors can show relationships to a degree allowed by the data. In this case, the data allowed four levels of similarity and a 'solution' expressed in a cluster analysis returned those levels. Classification schemes seek neatness and order. That a system based upon judgements from subjects will normally lack neatness casts suspicion on neat schemes. Henning's odor prism, which he claimed to have erected from empirical results, had the same neatness as the Zwaardemaker–Linnaean scheme. As the experimentalist Eleanor Gamble (Gamble, 1916) noted: 'Its very neatness is against it' (p. 137). Investigations bore out her suspicions.

categorically alike or because they come from seasonings (Chastrette *et al.*, 1988)? The stimuli impose some constraints, but much of what subjects express depends upon their internally generated strategies.

Once knowledgeable about what the person 'contributes' to the classification, one cannot view a classification scheme as a rendition of truth, but of someone's rendition of truth. As Hendrik Zwaardemaker (Zwaardemaker, 1925), the first major olfactory physiologist and the author of a well-known classification scheme, wrote: 'If we choose to give the same name to several similar odors, it could often happen that the evidence of our designation eludes other people . . . and one may question the legitimacy of our identification' (p. 179). In fact, others did question the legitimacy of Zwaardemaker's scheme (Titchener, 1912; Henning, 1916; Hazzard, 1930), as one person will inevitably question another's subjective organization of almost any complex matter. That, though, is the crux. As noted above, the measurements of interest should not be negotiable.

The notion that one scheme is right and another wrong is implicit in disputes over the 'validity' of one scheme versus another. History does not endorse any particular scheme as more valid than the rest, though some have proven quite useful for other reasons. For example, the sixfold scheme of Hans Henning (Henning, 1916) did not survive empirical scrutiny (Gamble, 1929; Boring, 1942; Harper *et al.*, 1968; Cain, 1978), but it did spur progress in the measurement of quality. The techniques of Henning's critics represented methodological advances, including odor profiling (Dimmick, 1922, 1927; Hazzard, 1930), triadic comparisons (Macdonald, 1922; Findley, 1924; Bentley, 1926) and direct ratings of similarity (Dimmick, 1927). John Amoore's (Amoore, 1962a,b, 1964) scheme of seven odor 'primaries', and the accompanying models of receptors based largely on structural least-common denominators within categories, received modest support at best (Döving, 1965; Köster, 1965; Amoore and Venstrom, 1967; Johnson, 1967; Amoore, 1970) [see also (Buck and Axel, 1991; Wang *et al.*, 1998)]. However, in the course of his work and that of others (Beets, 1978), it became evident that quality does relate to chemical structure, though specifics have proven elusive. Hypotheses regarding determinants of quality can certainly direct research, though there is something rather too facile and seductive about the search for least-common denominators with forgiving outcome variables (Rossiter, 1996).

Sorting

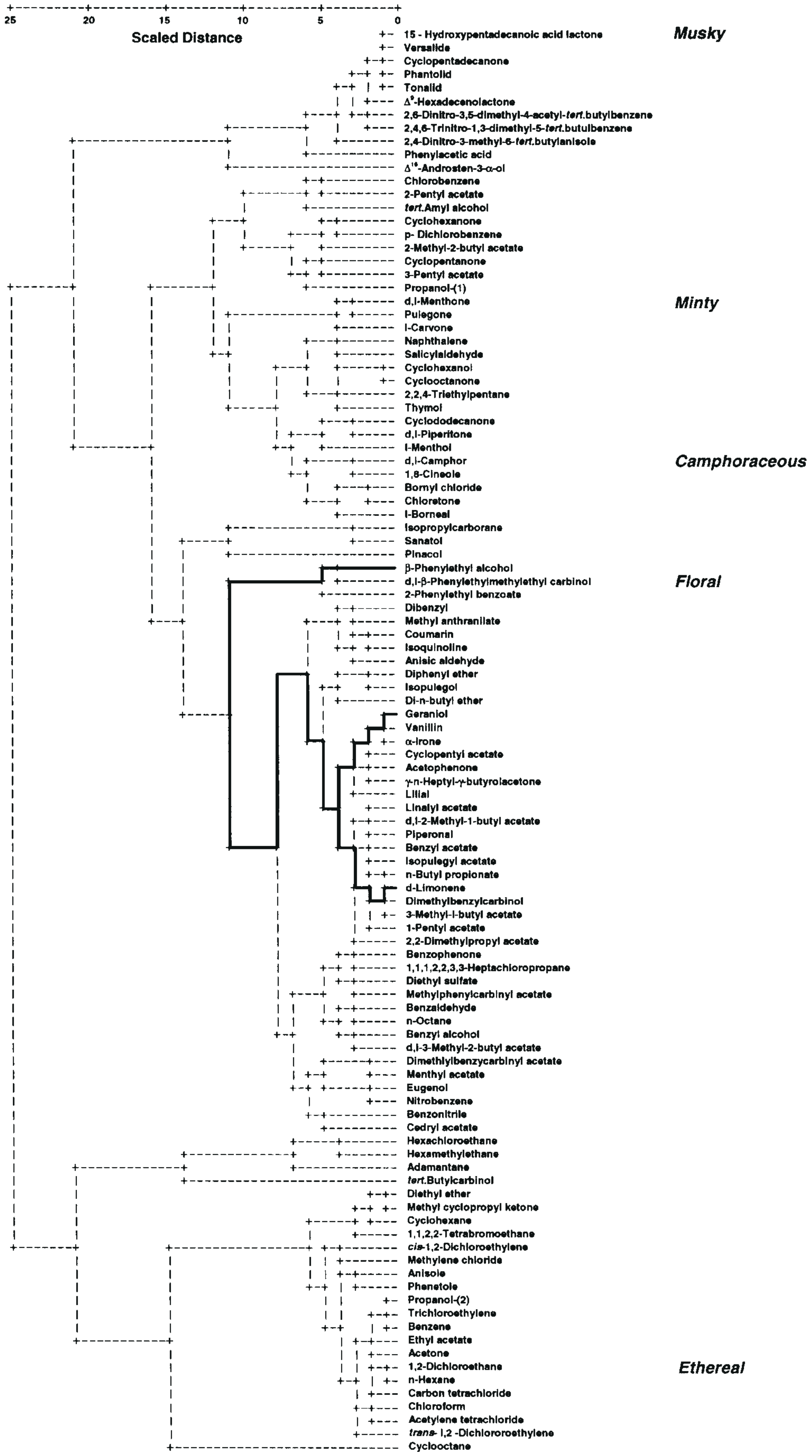
Some investigators have asked subjects to sort odors by qualitative resemblance (Lawless, 1989; MacRae *et al.*, 1990, 1992; Stevens and O'Connell, 1996). When a subject places two odors in the same group it indicates similarity. The technique has naive subjects do in a controlled setting what classifiers such as Linnaeus and Zwaardemaker did by personal observation. Lay subjects, however, generally have no knowledge of the composition or origin of the substances they evaluate. Could these subjects in their ignorance of origin offer more objectivity than experts? Not likely. No matter who does it, sorting still qualifies as subjective. Average data show reasonable test–retest reliability (MacRae *et al.*, 1990, 1992), but subjects differ in the number of categories they form. Sorting might in principle reflect true individual differences in reception, but even subjects within special subgroups differ in their categories (Stevens and O'Connell, 1996). Subjects accordingly seem to show diversity in their *decisions* about which odors belong in the same group. An attempt to force consistency by setting the number of categories simply imposes arbitrary structure on the data (Lawless, 1989).

Profiling

Profiling of odors relies on ratings of the applicability of odor-related words or reference odorants for their relevance to a test odorant. The words usually comprise fixed descriptors (Pilgrim and Schutz, 1957; Yoshida, 1964b; Harper *et al.*, 1968; Moskowitz and Barbe, 1977; Coxon *et al.*, 1978; Dravnieks *et al.*, 1978; Dravnieks, 1982; Laing and Willcox, 1983). The use of a list imposes some constraint on outcome and therefore plays it own role (see Figure 1). In the widely known scheme of Dravnieks (Dravnieks, 1985), subjects rate the applicability of each of 146 descriptors (e.g. cardboard-like, fishy, leather-like, sauerkraut-like) to a given test-odor on a scale from 0 to 5. In the other type of profiling, subjects evaluate odors against chemical references (Crocker, 1945; Schutz, 1964; Wright and Michels, 1964; Amoore and Venstrom, 1967; Yoshida, 1975; Polak *et al.*, 1978). Wright and Michel (Wright and Michel, 1964), for example, gave their subjects a scale of 1 to 7 to indicate the qualitative similarity between a test-odor and nine references. Amoore (Amoore, 1969) gave his subjects a scale of 1 to 9 to indicate similarity between 107 test odors and seven references (Figure 5).

Dravnieks's list of 146 descriptors covers a broad range of qualities, but Dravnieks himself anticipated that investigators would need to add more depending on the stimuli under scrutiny. Specialists have tailored systems to the specifications of the substances they evaluate. Noble's 'wine aroma wheel', for example, includes 87 notes relevant to wine (Noble *et al.*, 1984, 1987). With the need for so many verbal descriptors to obtain resolution of quality, one wonders how many chemical references it would take to characterize a given odor. Hundreds, as with words? Certainly more than seven or nine. Without a comprehensive list of descriptors (or references), one risks obtaining only a parochial picture of similarity. To illustrate, suppose one wished to profile a group of esters present in various fruits. If the list failed to include an adequate number of fruit-relevant descriptors (e.g. pineapple, banana, orange), one might underestimate differences among the esters. On the other hand, if one tailored the scheme to this group of chemicals by including only fruity descriptors, one might overestimate their differences compared with other stimuli. To a considerable extent, therefore, the experimenter's choice of descriptors determines outcome.

Almost needless to say, the use of verbal descriptors assumes subjects to have a common base of olfactory experience and to use the same words in the same way to describe sensations. That subjects differ both in the kind and number of descriptors they apply to a given odor makes this assumption questionable (Dravnieks *et al.*, 1978; Dravnieks, 1982, 1985; Cain *et al.*, 1998). Individual differences in reception accordingly become confounded with individual differences in cognition, culture and experience (Wysocki *et al.*, 1991). Dravnieks chose to average across 15 or more subjects and to purge aggregate profiles of infrequently used descriptors. Average profiles created in this fashion exhibited impressive reliability, but of course provided no information about individual differences (Dravnieks *et al.*, 1978; Dravnieks, 1982, 1985; Jeltema and Southwick, 1986); nor do they have a counterpart in animal research or offer a simple distillation of the profile into an index.


25
20
15
10
5
0
Scaled Distance
Musky
Minty
Camphoraceous
Floral
Ethereal
15 - Hydroxypentadecanoic acid lactone
Versalide
Cyclopentadecanone
Phantolid
Tonalid
Δ^9-Hexadecenolactone
2,6-Dinitro-3,5-dimethyl-4-acetyl-*tert*.butylbenzene
2,4,6-Trinitro-1,3-dimethyl-5-*tert*.butulbenzene
2,4-Dinitro-3-methyl-6-*tert*.butylanisole
Phenylacetic acid
Δ^{16}-Androsten-3-α-ol
Chlorobenzene
2-Pentyl acetate
tert.Amyl alcohol
Cyclohexanone
p- Dichlorobenzene
2-Methyl-2-butyl acetate
Cyclopentanone
3-Pentyl acetate
Propanol-(1)
d,l-Menthone
Pulegone
l-Carvone
Naphthalene
Salicylaldehyde
Cyclohexanol
Cyclooctanone
2,2,4-Triethylpentane
Thymol
Cyclododecanone
d,l-Piperitone
l-Menthol
d,l-Camphor
1,8-Cineole
Bornyl chloride
Chloretone
l-Borneal
Isopropylcarborane
Sanatol
Pinacol
β-Phenylethyl alcohol
d,l-β-Phenylethylmethylethyl carbinol
2-Phenylethyl benzoate
Dibenzyl
Methyl anthranilate
Coumarin
Isoquinoline
Anisic aldehyde
Diphenyl ether
Isopulegol
Di-n-butyl ether
Geraniol
Vanillin
α-Irone
Cyclopentyl acetate
Acetophenone
γ-n-Heptyl-γ-butyrolacetone
Lilial
Linalyl acetate
d,l-2-Methyl-1-butyl acetate
Piperonal
Benzyl acetate
Isopulegyl acetate
n-Butyl propionate
d-Limonene
Dimethylbenzylcarbinol
3-Methyl-l-butyl acetate
1-Pentyl acetate
2,2-Dimethylpropyl acetate
Benzophenone
1,1,1,2,2,3,3-Heptachloropropane
Diethyl sulfate
Methylphenylcarbinyl acetate
Benzaldehyde
n-Octane
Benzyl alcohol
d,l-3-Methyl-2-butyl acetate
Dimethlylbenzycarbinyl acetate
Menthyl acetate
Eugenol
Nitrobenzene
Benzonitrile
Cedryl acetate
Hexachloroethane
Hexamethylethane
Adamantane
tert.Butylcarbinol
Diethyl ether
Methyl cyclopropyl ketone
Cyclohexane
1,1,2,2-Tetrabromoethane
cis-1,2-Dichloroethylene
Methylene chloride
Anisole
Phenetole
Propanol-(2)
Trichloroethylene
Benzene
Ethyl acetate
Acetone
1,2-Dichloroethane
n-Hexane
Carbon tetrachloride
Chloroform
Acetylene tetrachloride
trans- l,2 -Dichlororoethylene
Cyclooctane

Direct ratings of similarity

In various investigations, subjects have been instructed to give numerical ratings of similarity to all pairwise combinations among sets of odors (Engen, 1964; Yoshida, 1964a,b; Döving and Lange, 1967; Woskow, 1968; Berglund *et al.*, 1973; Dravnieks, 1974; Moskowitz and Gerbers, 1974; Davis, 1979; Schiffman, 1981, 1984; Seeman *et al.*, 1989). Unlike profiles, direct ratings do not require the experimenter to derive a measure of similarity indirectly from either words or other odors. The numerical judgement stands for the similarity.

Direct scaling of similarity has focused in some measure on data for individuals to address whether subjects rate similarity similarly. Most studies showed poor agreement among individuals (Dimmick, 1927; Yoshida, 1964a,b; Gregson, 1972; Berglund *et al.*, 1973). Intersubject variation is not necessarily random, as Dimmick (Dimmick, 1927) first noted when his four subjects separated themselves into two pairs, with substantial agreement within pairs but poor agreement between them. Much later, Gregson (Gregson, 1972) factor-analyzed a matrix of inter-correlations between subjects and found several factors. Different subjects showed strong loadings on different factors (Davis, 1979). It appears that different subjects may judge similarity by different criteria.

Average ratings, which essentially treat differences between subjects as noise, seem to offer somewhat degenerate information regarding similarity. Multidimensional analyses of average ratings typically yield a dominant hedonic dimension/factor, with one or two others that defy clear interpretation (Yoshida, 1964a,b, 1972; Woskow, 1968; Moskowitz and Gerbers, 1974; Schiffman, 1981, 1984). Faced with an unfamiliar task and lacking clear instructions regarding criteria, subjects may resort to judgements of pleasantness as a way to convey some information. Beyond the salient aspect of pleasantness, the task may be analogous to asking someone to rate the similarity of a pencil and a baseball bat. One person may judge on the basis of size and find the objects dissimilar, whereas another may judge on the basis of shape and find them similar. Even if individuals use personal criteria consistently, the differences in non-hedonic criteria may tend to cancel out in the averaging. The hedonic dimension may survive because it alone figures in the ratings of most subjects.

Some investigators have worked with related chemicals that differ from one another in regular and specific ways (Engen, 1964; Southwick and Schiffman, 1980; Schiffman and Leffingwell, 1981). In such cases, subjects show better agreement, perhaps because related odors do not differ along so many dimensions and because relevant criteria may define themselves.

Data from different laboratories agree, though imperfectly [see (Callegari *et al.*, 1997) for a review of studies with some common test-odors]. Some inconsistencies could have come from methodological differences, but since different studies employed different sets of odors, some differences could have come from effects of context. Kurtz and collaborators (Kurtz *et al.*, 2000) found ratings of similarity among a group of odorants changed with substitution of one odor in the group. Thus, subjects do change the criteria by which they judge similarity.

Numerical ratings, such as those used in direct scaling of similarity, occupy a very important place in psychophysical research. They have generated much useful data, mostly on perceived intensity. Some investigators have maintained that ratings have high validity and some that they have virtually none. Some have argued for the validity of just one class of rating, e.g. magnitude estimation and related techniques, and some just for another kind, e.g. category rating.

Anderson (Anderson, 1982) has set out criteria to establish the linearity of numerical ratings with the internal representations of stimuli. The criteria entail collection of data on additivity—a matter not always possible to do—but the spirit of such tests has considerable appeal. If, for example, one could establish lawfulness of similarity estimates for the perceived quality of mixtures where an

Figure 5 Where to make the cut? A dendrogram computed by the present authors from judgements of similarity, in this case of 107 test odors against seven reference odors (Amoore, 1969), illustrates that an empirically based scheme will display more complexity than a classical classification scheme. A search for categories, or primaries in Amoore's terms, will depend upon the odors used as test stimuli and those used as references (in this case: 15-hydroxypentadecanoic acid lactone for the musk primary, D,L-menthone for the minty primary, 1,8-cineole for the camphoraceous primary, D,L-β-phenylmethylethyl carbinol for the floral primary and 1,2-dichloroethane for the ethereal primary, formic acid for pungent and dimethyl disulfide for putrid), the alphabet of the description. Although Amoore had reference odors for pungent and putrid, he had no test odors with those predominant qualities. Because Amoore oversampled musks, so designated because of their odors rather than their structure or functionality, he obtained the rather distinct musky cluster seen at the top of the tree. (The names of the primary categories stand opposite their respective references.) At the level of dissimilarity where the musks break away from the rest of the set, the camphoraceous and minty odors had not yet differentiated themselves into their respective clusters, nor had the floral cluster revealed its diversity, with spicy and floral subclusters. Thiboud (Thiboud, 1991) placed geraniol in the rosy class of odors (Figure 1), though noted its fruitiness. Zwaardemaker (Zwaardemaker, 1895) placed geraniol among the citrous odors, not among florals. The dendrogram illustrates that with their seven reference odors, five of consequence, Amoore's subjects found geraniol more like the citrous odor of D-limonene than the floral (rosy) odor of phenyl ethyl alcohol, as Zwaardemaker's scheme implied. With different references, however, Amoore's subjects may have given a different answer (Callegari *et al.*, 1997). In fairness to Amoore, the judgements of similarity he collected were meant to test hypotheses about odor quality and stereochemical properties. He did not necessarily intend the data to serve an archival purpose and presumably they will not. Indeed, no data obtained with just a few references, either odors or words, would serve well archivally. On simple perceptual grounds, one could use these data to support various possible schemes of the perceptual organization of odors, none probably any more valid than another. The scheme one ended up with would depend on the cut through the dendrogram.

experimenter could exert some control of quality, then standard means of testing for the additivity of estimates of similarity might qualify them in a way that no one has done for any other of the methodologies. This would not make their outcome isometric with animal data, nor would it necessarily prove that such estimations would have the requisite sensitivity, but the possibility that methodology to measure similarity could earn its stripes should remain open.

Triadic comparisons

In triadic comparisons, subjects pick the most similar and the least similar pairs of odors within triads (MacRae *et al.*, 1990, 1992). In this technique, subjects need not quantify similarity, but need make only ordinal judgements. As with ratings of similarity, different subjects might use different criteria. To our knowledge, no recent studies have examined individual differences. However, researchers who used similar techniques to evaluate Henning's prism (Macdonald, 1922; Findley, 1924; Bentley, 1926) reported that subjects differed widely in their judgements and suggested that subjects do in fact employ different criteria. The matter remains open.

Measures of performance

For more than a century, experimental psychologists have grappled with the issue of whether to build psychological science on studies of the content of the mind or on studies of human abilities (Cattell, 1893). To a large degree, one's position on the matter dictated permissible methodologies. If in the former camp, a report of mental content, such as a direct rating of the brightness of a scene, was permissible. If in the latter camp, quantification of perceived brightness would need to be approached less mentalistically. One could ask a subject to match one brightness to another (and by enumerating errors over trials could measure accuracy), or to state whether one brightness exceeded another, or otherwise to discriminate brightnesses, but not to quantify them directly. Why not? Because differences in ratings could have origins outside perceived brightnesses. The same issue exists here.

The rise of cognitive psychology and cognitive science has seen increasing interest in mental content, though cognitive scientists would find the term 'mental operations' more agreeable, but with refinement of experimental techniques to deduce operations via tests of ability, e.g. how fast subjects might notice features of complex stimuli. In this way, they avoid the pitfalls of 'mentalism'. The same could in principle occur in the investigation of odor quality. In olfactory psychophysics, investigators have examined how fast and how accurately subjects can sort odor mixtures on quality in order to determine if subjects can selectively attend to one qualitative attribute and ignore another. Whether or not subjects can do this for all odors, some odors or none has implications for olfactory coding (Schwartz *et al.*, 1987). Whether subjects can 'find' the correct number of constituents in a mixture may also have such relevance (Livermore and Laing, 1998). As unlikely as it may seem on the surface, these tests of ability, and various others (e.g. capacity to identify odors), could serve to examine aspects of structure–activity relationships. There are no *a priori* limits on the number or variety of tests of ability that one might use to address aspects of coding or structure–activity relationships.

Cross-adaptation

Exposure to an adapting-odor tends to raise the threshold for, lower the perceived intensity of and increase simple reaction time to a test-odor (Köster and de Wijk, 1991; Cometto-Muñiz and Cain, 1995). In general, the largest effects occur when adapting- and test-odors nominally match (self-adaptation), but adaptation can also occur when they differ (cross-adaptation). Scholars have proposed that cross-adaptation occurs when two odors stimulate overlapping groups of sensory channels or physiological mechanisms (Zwaademaker, 1925; Cheesman and Mayne, 1953; Moncrief, 1956; Cain and Polak, 1992). Could this functional similarity lead to an index of perceived similarity? Cross-adaptation, at least as defined as changes in threshold and reaction time, could qualify as reasonably objective.

Cross-adaptation could yield the desired index if degree of cross-adaptation reflects perceived similarity. Some studies have produced evidence consistent with this notion (e.g. Cheesman and Mayne, 1953; Moncrief, 1956; Todrank *et al.*, 1991; Cain and Polak, 1992). However, phenomenologically similar odors do not always cross-adapt and dissimilar ones sometimes do (O'Connell *et al.*, 1994; Pierce *et al.*, 1995). Further, asymmetric or non-reciprocal cross-adaptation often occurs (Cain and Engen, 1969; Cain, 1970; Köster, 1971; deWijk, 1989; Todrank *et al.*, 1991; Stevens and O'Connell, 1996). Both results suggest that degree of cross-adaptation entails more than sensory similarity. Finally, exposure to one odorant can sometimes *enhance* sensitivity to another odorant (Engen and Bosack, 1969; Berglund *et al.*, 1971; de Wijk, 1989; Stevens and O'Connell, 1996).

A valid scale of similarity may help investigators understand the complexities of cross-adaptation. In this context, it might prove instructive to compare degree of cross-adaptation with both perceived similarity and similarity of molecular properties (Pierce *et al.*, 1995). This approach, combined with physiological studies of receptors and central neural mechanisms, might shed light on the mechanisms of adaptation and the coding of odor quality. However, investigators have yet to demonstrate a clear relationship between degree of cross-adaptation and perceived similarity.

Adaptation could potentially play a role whenever subjects must make repeated judgements within a session (this caution also applies to the discriminative techniques discussed below). Repeated exposure tends to attenuate intensity, but might also decrease subjects' ability to discern differences in quality independent of changes in intensity. No empirical evidence exists to support this notion. Subjectively, however, the phenomenon can prove quite compelling. For example, those who work in retail sales of fragrance often caution against sampling too many perfumes in rapid succession, lest they begin to smell alike. Future research could examine the effects of repeated exposure, both to gain a better understanding of olfactory processing and to determine an inter-trial interval that would ensure a reasonable degree of independence between judgements.

Discrimination

At the most fundamental level, a sensory system does two things: it responds to a certain type of energy (radiative, mechanical, chemical) and within that type it resolves diffeences in kind (wavelength, molecular properties). Discrimination measures success of resolution. Success at resolution across many stimuli holds a key to what aspects of the stimulus are important. Beneath the surface, almost all tests of olfactory performance prove to be limited in some measure by the property of discrimination.

In the most basic discriminative paradigm, subjects receive some pairs of odors that do and some that do not differ in quality (Doty, 1991). In other paradigms, subjects receive three odors during a trial (Jones and Elliot, 1975; Eskenazi *et al.*, 1983, 1986; Hormann and Cowart, 1993). In an oddity paradigm, two of the three odors match and subjects must identify the 'odd' one. In an ABX paradigm, subjects must chose which of two different 'comparison-odors' matches a standard. In all cases, the frequency of incorrect responses represents qualitative similarity provided the odors do not differ in perceived intensity or trigeminal impact (such could serve as cues for discrimination).

Techniques of discrimination minimize or eliminate subjectivity. Accordingly, discriminative results are essentially non-negotiable. Indeed, there would be little disagreement about most results based on discrimination. We are not naive enough to think that data obtained by discrimination can invariably be deposited in some vault, but this is probably more true of such data than of estimates of similarity that purport to indicate how similar rose odor is to peanut odor. Hence, if data on discriminability exist, they could in principle be dovetailed to other discrimination-based data without harm. As tests of performance, rather than tests of mental content, they can have exact counterparts in animal studies. Recent similarities in outcome of discriminations between squirrel monkeys and humans reinforce the virtue of such comparability (Laska and Hudson, 1993; Laska and Freyer, 1997; Laska *et al.*, 1999).

Techniques of discrimination can resolve small differences in quality, even for individuals (Laska and Freyer, 1997; Laska and Teubner, 1999a; Laska *et al.*, 1999). There is no *a priori* limitation on the number of judgements an individual can contribute to a discriminative result. The data from an individual could in principle be collected with enough precision to distinguish heretofore unexplored subtleties, e.g. those present within family members versus those outside the family. Indeed, if one wanted to know whether another technique allowed for good resolution, one would use a direct discriminative technique as a standard.

The sensitivity of discrimination, an asset for resolution of small differences, becomes a liability if odors differ markedly. Performance in simple discrimination may then lie at asymptotic level. Only a few investigators have used discriminability to assess similarity outside studies of enantiomers, where the issue was whether the odors differed more than how much they differed (Theimer and McDaniel, 1970; Friedman and Miller, 1971; Jones and Elliot, 1975; Hummel *et al.*, 1992; Hormann and Cowart, 1993).

On general grounds, if forced to choose between adequate measurement of small differences versus large differences, one should choose the former. Various investigators have compared a wide variety of odorants in hope of capturing the major dimensions of 'odor space' (Schutz, 1964; Woskow, 1968; Yoshida, 1975; Schiffman, 1981). Accordingly, the researchers compared molecules that differed considerably in molecular parameters. This approach assumes that the many candidates for appropriate molecular parameters appear in a breadth necessary to see significant relationships. One might need to study hundreds or thousands of stimuli to uncover the relevant molecular parameters. Mapping changes that occur with gradual, systematic changes in molecular parameters would seem a shorter path to the goal of understanding. Since discrimination resolves small differences in quality, it seems well-suited to this task.

Laska and colleagues have begun to explore the discriminability of substances that differ by methylene groups in aliphatic series (Laska and Freyer, 1997; Laska and Teubner, 1999b; Laska *et al.*, 1999). Within a series, discriminability increased with difference in carbon chain-length, though did reach asymptotic levels at a difference beyond about three carbon atoms. It would be desirable to map trends in quality over a wider range, an entirely achievable goal, via: (i) addition of a new dependent variable to the task of discrimination, without any change in the stimuli themselves; or (ii) increasing the difficulty of the discrimination with a change in the stimuli

A large literature suggests that the time needed to compare two physical stimuli varies inversely with the perceptual distance between them, even beyond the point of perfect discriminability (Cattell, 1902; Henmon, 1906; Kellogg, 1931; Woodworth and Schlosberg, 1954; Crossman, 1955; Welford, 1960; Vickers, 1980). Wise and Cain (Wise and Cain, 2000) found that measures of discriminability based on latency to discriminate correlated strongly with measures based on errors of discrimination, but latency-based

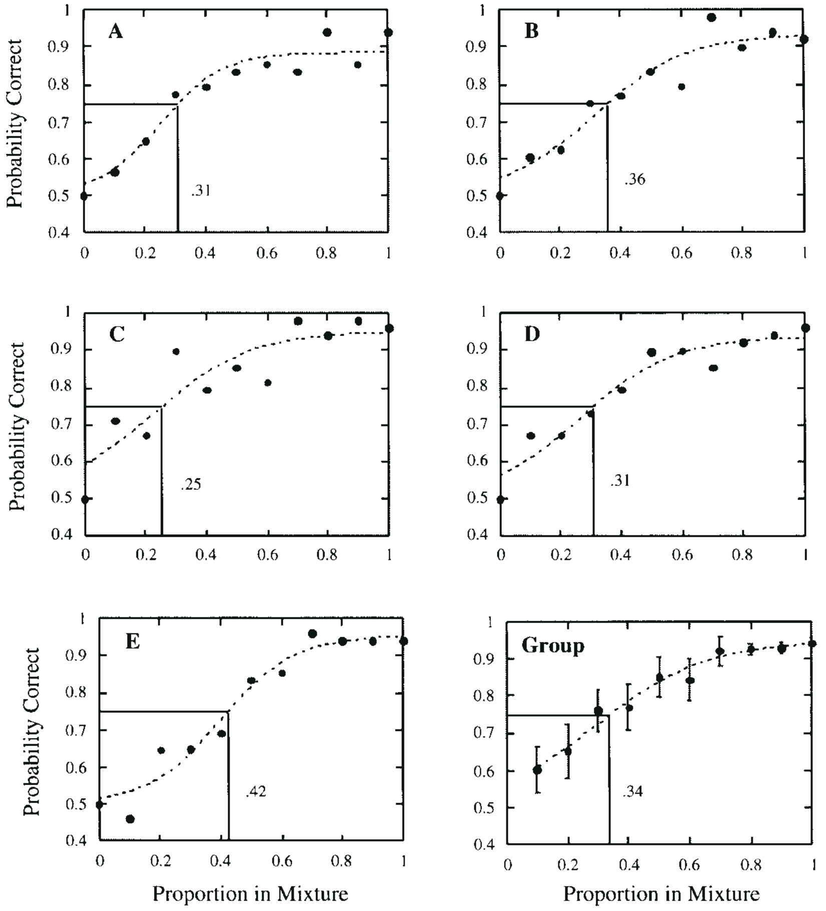


Figure 6 Psychometric functions for the discriminability of citral from eugenol and vice versa as the proportion of one or the other odor increased from zero to 1.0 (Olsson and Cain, 2000). Each of the four subjects made 96 judgements per point. Chance performance equaled 0.5. The line dropped to the abscissa represents the differential threshold (0.75 correct) for quality discrimination. In principle, an odor could be fixed in an *n*-dimensional space by measurement of its discriminability from various others. Odors with very different qualities would presumably require less dilution for criterion discriminability than odors with very similar qualities, although this would need to be verified empirically.

measures provided better resolution among odor-pairs. Application of latency-to-discriminate to mixtures endorsed the conclusion that the odor qualities of binary combinations lie approximately midway between the qualities of their components. More importantly, the discriminabilities between pairs of mixtures not previously measured proved predictable from the discriminabilities of their components. Latency to discriminate shows considerable promise as a dependent variable. It has begun to earn its stripes.

In the perceptual 'space' between any two odors, one can conceive of a continuum over which one odor becomes gradually transformed into the other by progressive dilution and concentration. Hence, as clove odor becomes increasing diluted with lemon odor its discriminability from clove grows progressively and its discriminability from lemon diminishes. In principle, discriminations between a binary mixture and an unmixed component should prove more difficult when the two components are similar, and less difficult when the two components are not similar. Hence, beginning with two similar odors, progressive dilution of one by the other should reflect itself in poor differential sensitivity (Figure 6) (Duncan *et al.*, 1992; Olsson and Cain, 2000). Variations such as this may need to earn their stripes from collection of data convergent with that of other techniques.

One might also increase the difficulty of discrimination by lowering the perceived intensity of the stimuli. Discriminations of hue and pitch become more difficult at lower intensities (Shower and Biddulph, 1931; Brown, 1951; Harris, 1952; Walraven and Bouman, 1966; Reitner *et al.*, 1992). This undoubtedly holds true for olfactory discriminations as well (Jones and Elliot, 1975; de Wijk and Cain, 1994). Odor quality may change with changes in concentration for some molecules (Moncrief, 1967; Gross-Isseroff and Lancet, 1988). Such effects could complicate matters to a certain extent from the standpoint of measurement, but might also shed light on olfactory coding. Future investigations could determine how the discriminative odor-space, i.e. both absolute and relative differences among odors, changes with alterations of perceived intensity.

As mentioned above, future investigations could also determine how the discriminative odor-spaces of individuals differ. To what degree do normal subjects vary? How many subjects will we need to study for a general picture of similarity? Effective measurement of differences in quality might require answers to these questions.

Of course, individual differences provide opportunities as well as challenges. For example, the reduced discriminative capacity of patients with damage to certain areas of the brain might shed light on the processing of odor quality in the normal brain (Eskenazi *et al.*, 1986; Martinez *et al.*, 1993). The discrimination of persons with a specific anosmia [poor sensitivity to a certain chemical despite otherwise normal olfactory acuity (Amoore, 1969, 1970; O'Connell *et al.*, 1989; Stevens and O'Connell, 1991, 1996)] might also prove enlightening, though investigators have yet to define specific anosmia precisely [see (O'Connell *et al.*, 1994) for a discussion], much less explain it.

When considered across the sense-modalities, no area of sensory measurement has received more attention than that of discrimination. The auditory psychophysics of simple and moderately complex stimuli, developed by techniques of discrimination, of matching and of numerical scaling, has succeeded so well as to have reached a point of diminishing returns. The theory of signal detectability (TSD) played a significant role in this accomplishment. TSD offers many opportunities to turn ordinal data of the sort generated in studies of performance into metric indices (MacMillan and Creelman, 1991; Gescheider, 1997). Some such indices have already seen use in olfaction (Cain, 1977; Rabin and Cain, 1989; Swets, 1986; Cain and Potts, 1996). Whereas research could conceivably lead to a metric measure for a descriptive technique, it already offers choices for techniques based upon performance.

Conclusions

Understanding the relationship between molecular properties and odor quality arguably poses the single most significant problem in olfaction. It seems curious that in recent times no psychophysical laboratory has dedicated itself to systematic pursuit of the problem. Essential understanding will require more than just psychophysical correlations between structure and activity, whether in humans or other animals. It will require knowledge of interactions between molecules and receptors. Neither pursuit alone, though, will solve the problem, but this need not inhibit any level of approach. Any given molecule will undoubtedly stimulate more than one type of receptor, perhaps even hundreds, with probabilities reminiscent of those of resonance structures. How such stimulation reflects itself in a neural code will require modeling of mechanisms of neural computation at virtually every stage of the olfactory nervous system. How the neural code for this dual discriminative–regulatory modality varies with hunger, thirst, sexual needs, arousal, mood, metabolism, state of health, adaptation, and the presence of other odoriferous and nonodoriferous stimuli will remain issues of importance. Curiously, ancillary issues such as these have received more attention in general than the primary issue of how to define the stimulus in the first order.

Lack of psychophysical direction plays its role in the languor of olfactory structure–activity relationships. Psychophysical studies have rarely built on one another. Data from one study rarely seem trustworthy enough for another investigator to use them. Hence, every study starts anew, as if the first of what might become a long string of studies, but in fact never does (Cain, 1978). How many areas of sensory science can say: 'We have no archival data. If you bring us a new item [in this case an odorant], we will likely have no idea what its threshold will be and no idea what its perceived quality will be'? Oh, yes, there are trivial cases, a simple mercaptan or ester or amine, where any chemist can guess that the stimulus will smell skunky, fruity or fishy–uriny, something chemists could do well over a century ago, but such trivia aside, the field languishes as much for a way to accumulate data as anything. This is not an essential problem, for it has a solution. It is an 'accidental' problem in the Aristotelian sense. Unscrutinized psychophysical methodology has contributed to it and more circumspect measures based upon performance should help solve it. In the pursuit of an answer, we feel responsive to the eminent chemosensory scientist Lloyd Beidler (Beidler, 1976), who noted: 'Rigor in defining the quantity and quality of the [sensory] response measured must be as good as that used in determining the physicochemical properties of the stimulus molecule if reliable correlations [between structure and activity] are to be expected' (p. 295). Beidler shared Alexander Graham Bell's priority.

We have focused principally on choice of a methodology, but we recognize that a methodology based on performance comes at the price of time. Automated presentation of stimuli could reduce the burden and sorely needs development. Devices that test human beings automatically could, with appropriate modification, also test animals. Development of an animal psychophysics of odor quality could greatly speed up accumulation of data, which might conceivably have as much relevance to human perception as human responses themselves. Finally, in our focus on the measurement of activity in structure–activity relationships, we may have seemed to underestimate complexities of the characterization of molecules. Molecules of odorant can be characterized in considerable detail, but the chemical challenge in structure–activity relationships is to decide both the relevant type of characterization and the details. As Chastrette (Chastrette, 1997) indicated, choices for type include: (i) global properties, including boiling point, molar volume and calculated shape; (ii) geometric variables, commonly distances within a molecule; (iii) electrostatic variables; and (iv) fragments and patterns as epitomized in hypothesized osomophoric groups and odortopes (Mori and Shepherd, 1994). To sort out the rules will require testing of hypotheses with both adequate measures of structure and adequate measures of activity. This will undoubtedly require collaboration of scientists on both sides of the structure–activity equation.

Acknowledgements

Supported by research grant 5 RO1 DC 00284 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. We would like to thank three anonymous reviewers and Dr Craig Warren for useful comments.

References

Amoore, J.E. (1962a) *The stereochemical theory of olfaction 1: Identification of the seven primary odours*. Proceedings of the Scientific Section, Toilet Goods Association, No. 37 suppl., 1–12.

Amoore, J.E. (1962b) *The stereochemical theory of olfaction 2: Elucidation of the stereochemical properties of the olfactory receptor sites*. Proceedings of the Scientific Section, Toilet Goods Association, No. 37 suppl., 13–23.

Amoore, J.E. (1964) *Current status of the steric theory of odor*. Ann. N.Y. Acad. Sci., 116, 457–476.

Amoore, J.E. (1969) *A plan to identify most of the primary odors*. In Pfaffmann, C. (ed.), Olfaction and Taste III. Rockefeller University Press, New York, pp. 158–171.

Amoore, J.E. (1970) Molecular Basis of Odor. Charles C. Thomas, Springfield, IL.

Amoore, J.E. and **Venstrom, D.** (1967) *Correlations between stereochemical assessments and organoleptic analysis of odorous compounds*. In Hayashi, T. (ed.), Olfaction and Taste II. Pergamon Press, Oxford, pp. 3–17.

Anderson, N.H. (1982) Methods of Information Integration Theory. Academic Press, New York.

Beare, J.I. (1906) Greek Theories of Elementary Cognition from Alcmaeon to Aristotle. Clarendon Press, Oxford.

Beets, M.G.J. (1978) Structure–Activity Relationships in Human Chemoreception. Applied Science Publisher.

Beidler, L.M. (1976) *Taste and smell: review of G. Benz (Ed.), Structure–Activity Relationships in Chemoreception*. Trends Biochem. Sci., 1, 295–296.

Bentley, M. (1926) *Qualitative resemblance among odors*. Psychol. Monogr., 35, 144–151.

Berglund, B., Berglund, U., Engen, T. and **Ekman, G.** (1973) *Multidimensional analysis of twenty-one odors*. Scand. J. Psychol., 14, 131–137.

Boring, E.G. (1942) Sensation and Perception in the History of Experimental Psychology. Irvington, New York.

Brown, W.R.J. (1951) *The influence of luminance level on visual sensitivity to colour differences*. J. Opt. Soc. Am., 41, 684–688.

Buck, L. (1996) *Information coding in the vertebtrate olfactory system*. Annu. Rev. Neurosci., 19, 517–544.

Buck, L. and **Axel, R.** (1991) *A novel multigene family may encode odorant receptors: a molecular basis for odor recognition*. Cell, 65, 175–187.

Cain, W.S. (1970) *Odor intensity after self-adaptation and cross-adaptation*. Percept. Psychophys., 7, 271–275.

Cain, W.S. (1977) *Differential sensitivity for smell: 'noise' at the nose*. Science, 195, 796–798.

Cain, W.S. (1978) *History of research on smell*. In Carterette, E.C. and Friedman, M.P. (eds), Handbook of Perception: Tasting and Smelling. Academic Press, New York, Vol. VIA, pp. 197–229.

Cain, W.S. (1988) *Olfaction*. In Atkinson, R.C., Herrnstein, R.J., Lindzey, G. and Luce, R.D. (eds), Stevens' Handbook of Experimental Psychology: Perception and Motivation. Wiley, New York, Vol. 1, pp. 409–459.

Cain, W.S. and **Engen, T.** (1969) *Olfactory adaptation and the scaling of odor intensity*. In Pfaffmann, C. (ed.), Olfaction and Taste. Rockefeller University Press, New York.

Cain, W.S. and **Polak, E.H.** (1992) *Olfactory adaptation as an aspect of odor similarity*. Chem. Senses, 17, 481–491.

Cain, W.S. and **Potts, B.C.** (1996) *Switch and bait: probing the discriminative basis of odor identification via recognition memory*. Chem. Senses, 21, 35–44.

Cain, W.S., de Wijk, R.A., Lulejian, C., Schiet, F. and **See, L.C.** (1998) *Odor identification: perceptual and semantic dimensions*. Chem. Senses, 23, 309–326.

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

Cattell, J.McK. (1893) *On errors of observation*. Am. J. Psychol., 5, 285–293.

Cattell, J.McK. (1902) *The time of perception as a measure of differences in intensity*. Philos. Stud., 19, 63–68.

Chastrette, M. (1997) *Trends in structure–odor relationships*. SAR QSAR Environ. Res., 6, 215–254.

Chastrette, M., Elmouaffek, A. and **Sauvegrain, P.** (1988) *A multidimensional statistical study of similarities between 74 notes used in perfumery*. Chem. Senses, 13, 295–305.

Cheesman, G.H. and **Mayne, S.** (1953) *The influence of adaptation on absolute threshold measurements of olfactory stimuli*. Quart. J. Exp. Psychol., 5, 22–30.

Cometto-Muñiz, E.C. and **Cain, W.S.** (1995) *Olfactory adaptation*. In Doty, R.L. (ed.), Handbook of Olfaction and Gustation. New York, Marcel Dekker, pp. 257–282.

Coxon, J.M., Gregson, A.M. and **Paddick, R.G.** (1978) *Multidimensional scaling of perceived odour of bicyclo [2.2.1] heptane, 1,7,7-Trimethylbicyclo [2.2.1] heptane and Cyclohexane derivatives*. Chem. Senses Flav., 3, 431–441.

Crocker, E.C. (1945) Flavor. McGraw-Hill, New York.

Crossman, E.R.F.W. (1955) *The measurement of discriminability*. Quart. J. Exp. Psychol., 7, 176–195.

Davis, R.G. (1979) *Olfactory perceptual space models compared by quantitative methods*. Chem. Senses Flav., 4, 21–33.

de Wijk, R.A. (1989) *Temporal factors in human olfactory perception*. Unpublished doctoral dissertation, University of Utrecht.

de Wijk, R.A. and **Cain, W.S.** (1994) *Odor quality: discrimination versus free and cued identification*. Percept. Psychophys., 56, 12–18.

Dimmick, F.L. (1922) *A note on Henning's smell series*. Am. J. Psychol., 33, 423–25.

Dimmick, F.L. (1927) *The investigation of the olfactory qualities*. Psychol. Rev., 34, 321–335.

Doty, R.L. (1991) *Psychophysical measurement of odor perception in* *humans*. In Laing, D.G., Doty, R.L. and Breipohl, W. (eds), The Human Sense of Smell. Springer-Verlag, Berlin, pp. 95–151.

Döving, K.B. (1965) *Studies on the responses of bulbar neurons of frog to different odour stimuli*. Rev. Laryngol. (Suppl.), 845–854.

Döving, K.B. and **Lange, A.L.** (1967) *Comparative studies of sensory relatedness of odors*. Scand. J. Psychol., 8, 47–51.

Dravnieks, A. (1974) *A building-block model for the characterization of odorant molecules and their odors*. Ann. N.Y. Acad. Sci., 237, 144–163.

Dravnieks, A. (1982) *Odor quality: semantically generated multi-dimensional profiles are stable*. Science, 218, 799–801.

Dravnieks, A. (1985) Atlas of Odor Character Profiles, Vol. 61. American Society for Testing and Materials, Philadelphia, PA.

Dravnieks, A., Bock, F.C., Tibbets, M. and **Ford, M.** (1978) *Comparison of odors directly and through odor profiling*. Chem. Senses, 3, 191–220.

Duncan, H.J., Beauchamp, G.K. and **Yamazaki, K.** (1992) *Assessing odor generalization in the rat: a sensitive technique*. Physiol. Behav., 52, 617–620.

Engen, T. (1964) *Psychophysical scaling of intensity and quality*. Ann. N.Y. Acad. Sci., 116, 504–516.

Eskenazi, B., Cain, W.S., Novelly, R.A. and **Friend, K.** (1983) *Olfactory functioning in temporal lobectomy patients*. Neuropsychologia, 21, 365–374.

Eskenazi, B., Cain, W.S. and **Friend, K.** (1986) *Olfactory functioning in temporal lobectomy patients*. Neuropsychologia, 24, 553–562.

Findley, A.E. (1924) *Further studies of Henning's system of olfactory qualities*. Am. J. Psychol., 35, 436–445.

Fráter, G. and **Lamparsky, D.** (1991) *Synthetic products*. In Müller, P.M. and Lamparsky, D. (eds), Perfumes: Art, Science, Technology. Kluwer Academic Publishers, Dordrecht, pp. 533–628.

Friedman, L. and **Miller, J.G.** (1971) *Odor incongruity and chirality*. Science, 172, 1044–1046.

Gamble, E.M. (1916) *Taste and smell*. Psychol. Bull., 13, 134–137.

Gamble, E.M. (1929) *Psychology of taste and smell: status as of 1929*. Psychol. Bull., 16, 566–569.

Gescheider, G.A. (1997) Psychophysics: The Fundamentals, 3rd edn. Lawrence Erlbaum Associates, Mahwah, NJ.

Gregson, R.M. (1972) *Odour similarities and their multidimensional metric representation*. Multivar. Behav. Res., 4, 165–175.

Gross-Isseroff, R. and **Lancet, D.** (1988) *Concentration-dependent changes of percieved odor quality*. Chem. Senses, 13, 191–204.

Harper, R., Bate Smith, E.C. and **Land, D.G.** (1968) Odour Description and Odour Classification. American Elsevier, New York.

Harris, J.D. (1952) *Pitch discrimination*. J. Acoust. Soc. Am., 24, 750–755.

Hazzard, F.W. (1930) *A descriptive account of odors*. J. Exp. Psychol., 13, 297–331.

Henmon, V.A.C. (1906) *The time of perception as a measure of differences in sensations*. Arch. Philos. Psychol. Scient. Methods, 8, 1–75.

Henning, H. (1916) Der Geruch. Barth, Leipzig.

Hormann, C.A. and **Cowart, B.J.** (1993) *Olfactory discrimination of carvone enantiomers*. Chem. Senses, 18, 13–21.

Hummel, T., Hummel, C., Pauli, E. and **Kobal, G.** (1992) *Olfactory discrimination of nicotine-enantiomers by smokers and non-smokers*. Chem. Senses, 17, 13–21.

Jeltema, M.A. and **Southwick, E.W.** (1986) *Evaluation and applications of odor profiling*. J. Sens. Stud., 1, 123–136.

Johnson, J.W. (1967) *Experiments on the specificities of human olfaction. Part 2*. In Hayashi, T. (ed.), Olfaction and Taste II, Pergamon Press, Oxford, pp. 19–50.

Jones, F.N. and **Elliot, D**. (1975) *Individual and substance differences in the discriminability of optical isomers*. Chem. Senses Flav., 1, 317–321.

Kellogg, W.N. (1931) *Time of judgment in psychometric measures*. Am. J. Psychol., 43, 65–86.

Köster, E.P. (1965) *Adaptation, recovery, and specificity of olfactory receptors*. Rev. Laryngol. (Suppl.), 880–894.

Köster, E.P. (1971) *Adaptation and cross-adaptation in olfaction*. Unpublished doctoral dissertation, University of Utrecht.

Köster, E.P. and **de Wijk, R.A.** (1991) *Olfactory Adaptation*. In Laing, D.G., Doty, R.L. and Breipohl, W. (eds), The Human Sense of Smell. Berlin, Springer-Verlag, pp. 199–215.

Kurtz, D.B., White, T.L. and **Hayes, M.** (2000) *The labeled dissimilarity scale: a metric of perceptual dissimilarity*. Percept. Psychophys., 62, 152–161.

Laing, D.G. and **Willcox, M.E.** (1983) *Perception of components in binary odour mixtures*. Chem. Senses, 7, 249–264.

Laska, M.L. and **Freyer, D.** (1997) *Olfactory discrimination ability for aliphatic esters in squirrel monkeys and humans*. Chem. Senses, 22, 457–465.

Laska, M. and **Hudson, R.** (1993) *Assessing olfactory performance in a new world primate,* Saimiri sciureus. Physiol. Behav., 53, 89–95.

Laska, M. and **Tuebner, P.** (1999a) *Olfactory discrimination ability of human subjects for ten pairs of enantiomers*. Chem. Senses, 24, 161–170.

Laska, M. and **Tuebner, P.** (1999b) *Olfactory discrimination ability for homologous series of aliphatic alcohols and aldehydes*. Chem. Senses, 24, 263–270.

Laska, M., Trolp, S. and **Tuebner, P.** (1999) *Odor structure–activity relationships compared in human and nonhuman primates*. Behav. Neurosci., 113, 998–1007.

Lawless, H.T. (1989) *Exploration of fragrance categories and ambiguous odors using multidimensional scaling and cluster analysis*. Chem. Senses, 14, 349–360.

Livermore, A. and **Laing, D.G.** (1998) *The influence of odor type on the discrimination and identification of multi-component mixtures*. Physiol. Behav., 65, 311–320.

Macdonald, M.K. (1922) *An experimental study of Henning's system of olfactory qualities*. Am. J. Psychol., 33, 321–35.

MacMillan, N.A. and **Creelman, D.C.** (1991) Detection Theory: A User's Guide. Cambridge University Press, Cambridge.

MacRae, A.W., Howgate, P. and **Geelhoed, E.** (1990) *Assessing the similarity of odours by sorting and by triadic comparison*. Chem. Senses, 15, 691–699.

MacRae, A.W., Rawcliffe, T., Howgate, P. and **Geelhoed, E.N.** (1992) *Patterns of odour similarity among carbonyls and their mixtures*. Chem. Senses, 17, 119–225.

Martinez, B.A., Cain, W.S., de Wijk, R.A., Spencer, D.D., Novelly, R.A. and **Sass, K.J.** (1993) *Olfactory functioning before and after temporal lobe resection for intractable seizures*. Neuropsychology, 7, 351–363.

Moncrief, R.W. (1956) *Olfactory adaptation and odour likeness*. J. Physiol., 133, 301–316.

Moncrief, R.W. (1967) The Chemical Senses. Leonard Hill, London.

Mori, K. and **Shepherd, G.M.** (1994) *Emerging principles of molecular signal processing by mitral/tufted cells in the olfactory bulb*. Semin. Cell Biol., 5, 65–74.

Moskowitz, H. R. and **Barbe, C.D.** (1977) *Profiling of odors components and their mixtures*. Sens. Process., 1, 212–226.

Moskowitz, H. R. and **Gerbers, C.L.** (1974) *Dimensional salience of odors*. Ann. N.Y. Acad. Sci., 237, 1–16.

Murphy, G.L. and **Medin, D.L.** (1985) *The role of theories in conceptual coherence*. Psychol. Rev., 92, 289–316.

Napolitano, E., Giovani, E., Centini, M., Anselmi, C. and **Pelosi, P.** (1994) *Effect of sulfur substitution on the floral odor of tetrahydropyranyl and tetrahydrofuranyl ethers*. J. Agric. Food Chem., 42, 1332–1334.

Napolitano, E., Giovani, E., Ceccarelli, N. and **Pelosi, P.** (1996) *Tertiary alcohols with earthy odor*. J. Agric. Food Chem., 44, 2806–2809.

Noble, A.C., Arnold, R.A., Masuda, B.M., Pecore, S.D., Schmidt, J.O. and **Stern, P.M.** (1984) *Research note: progress towards a standardized system of wine aroma terminology*. Am. J. Enol. Viticult., 35, 107–109.

Noble, A.C., Buechsensenstein, J., Leach, E.J., Schmidt, J.O. and **Stern, P.M.** (1987) *Modification of a standardized system of wine aroma terminology*. Am. J. Enol. Viticult., 38, 143–146.

O'Connell, R.J., Stevens, D.A., Akers, R.P., Coppola, D.M. and **Grant, A.J.** (1989) *Individual differences in the quantitative and qualitative responses of human subjects to various odors*. Chem. Senses, 14, 293–302.

O'Connell, R.J., Stevens, D.A. and **Zogby, L.M.** (1994) *Individual differences in the perceived intensity and quality of specific odors following self- and cross-adaptation*. Chem. Senses, 19, 197–208.

Ohloff, G., Winter, B. and **Fehr, C.** (1991) *Chemical classification and structure–odour relationships*. In Müller, P.M. and Lamparsky, D. (eds), Perfumes: Art, Science, and Technology. Kluwer Academic Publishers, Dordrecht, pp. 287–330.

Olsson, M.J. and **Cain, W.S.** (2000) *Psychometrics of odor quality discrimination: method for threshold determination*. Chem. Senses, in press.

Pierce, J.D., Zeng, X.N., Aronov, E.V., Preti, G. and **Wysocki, C.J.** (1995) *Cross-adaptation of sweaty-smelling 3-methyl-2-hexenoic acid by a structurally-similar, pleasant-smelling odorant*. Chem. Senses, 20, 401–411.

Pilgrim, F.J. and **Schutz, H.G.** (1957) *Measurement of quantitative and qualitative attributes of flavor*. In National Academy of Sciences–National Research Council Symposium, 'Chemistry of Food Flavors'. National Academy Press, Washington, DC, pp. 47–58.

Polak, E., Trotier, D. and **Baliguet, E.** (1978) *Odor similarities in structurally related odorants*. Chem. Senses Flav., 3, 369–380.

Rabin, M.D. and **Cain, W.S.** (1989) *Attention and learning in the perception of odor mixtures*. In Laing, D.H., Cain, W.S., McBride, R.L. and Ache, B.W. (eds), Perception of Complex Smells and Tastes. Academic Press, Sydney, pp. 173–188.

Reitner, A., Sharpe, L.T. and **Zrenner, E.** (1992) *Wavelength discrimination as a function of intensity, duration, and size*. Vision Res., 32, 179–185.

Rossiter, K.J. (1996) *Structure–odor relationships*. Chem. Rev., 96, 3201–3240.

Schiffman, S.S. (1981) *Characterization of odor quality utilizing multidimensional scaling techniques*. In Moskowitz, H.R. and Warren, C.B. (eds), Odor Quality and Chemical Structure. American Chemical Society, Washington, DC, pp. 1–19.

Schiffman, S.S. (1984) *Mathematical approaches for quantitative design of odorants and tastants*. In Warren, C. and Walradt, J. (eds), Computers in Flavor and Fragrance Research. American Chemical Society, Washington, DC, pp. 33–50.

Schiffman, S.S. and **Leffingwell, J.C.** (1981) *Perception of odors of simple pyrazines by young and elderly subjects: multidimensional analysis.* Pharmacol. Biochem. Behav., 14, 787–798.

Schutz, H.G. (1964) *A matching-standard method for characterizing odor qualities*. Ann. N.Y. Acad. Sci., 116, 517–526.

Schwartz, A.L., Rabin, M.D. and **Cain, W.S.** (1987) *Perceptual separability and integrality in odor discrimination*. Ann. NY Acad. Sci., 510, 595–596.

Seeman, J.I., Ennis, D.M., Secor, H.V., Clawson, L. and **Palen, J.** (1989) *The perceptual similarity of substituted benzenes and pyridines as a function of steric hindrance*. Chem. Senses, 14, 395–405.

Shower, E.G. and **Biddulph, R.** (1931) *Differential pitch sensitivity of the ear*. J. Acoust. Soc. Am., 3, 275–287.

Smith, E.E. and **Medin, D.L.** (1981) Categories and Concepts. Harvard University Press, Cambridge, MA.

Southwick, E. and **Schiffman, S.S.** (1980) *Odor quality of pyridyl ketones*. Chem. Senses, 5, 343–357.

Stevens, D.A. and **O'Connell, R.J.** (1991) *Individual differences in thresholds and quality reports of human subjects to various odors*. Chem. Senses, 16, 57–67.

Stevens, D.A. and **O'Connell, R.J.** (1996) *Semantic-free scaling of odor quality*. Physiol. Behav., 60, 211–215.

Stevens, S.S. (1950) *Mathematics, measurement, and psychophysics*. In Stevens, S.S. (ed.), Handbook of Experimental Psychology. Wiley, New York, pp. 1–49.

Stevens, S.S. (1960) *The psychophysics of sensory function*. Am. Scient., 48, 226–253.

Stone, H. and **Sidel, J.L.** (1985) Sensory Evaluation Practices. Academic Press, Orlando, FL.

Swan, J.S. and **Burtles, S.M.** (1980) *Quality control of flavour by the use of integrated sensory analytical methods at various stages of Scotch whisky production*. In van der Starre, H. (ed.), Olfaction and Taste VII. IRL Press, London, pp. 451–452.

Swets, J.A. (1986) *Form of empirical ROC's in discrimination and diagnostic tasks: implications for theory and measurement of performance*. Psychol. Bull., 99, 181–198.

Theimer, E.T. and **McDaniel, M.R.** (1970) *Odor and optical activity*. J. Soc. Cosmet. Chem., 22, 15–26.

Thiboud, M. (1991) *Empirical classification of odours*. In Müller, P.M. and Lamparsky, D. (eds), Perfumes: Art, Science, and Technology. Kluwer Academic Publishers, Dordrecht, pp. 253–286.

Titchener, E.B. (1912) A Textbook of Psychology. MacMillan, New York.

Todrank, J., Wysocki, C.J. and **Beauchamp, G.K.** (1991) *The effects of adaptation on the perception of similar and dissimilar odors*. Chem. Senses, 16, 467–482.

Vickers, D. (1980) *Discrimination*. In Welford, A.T. (ed.), Reaction Times, Academic Press, London, pp. 25–72.

Walraven, P.L. and **Bouman, M.A.** (1966) *Fluctuation theory of colour discrimination of normal trichromats*. Vision Res., 6, 567–586.

Wang, F., Nemes, A., Mendelsohn, M. and **Axel, R.** (1998) *Odorant receptors govern the formation of a precise topographic map.* Cell, 93, 47–60.

Welford, A.T. (1960) *The measurement of sensory-motor performance: survey and re-appraisal of twelve years' progress*. Ergonomics, 3, 189–230.

Wise, P.M. and **Cain, W.S.** (2000) *Latency and accuracy of discriminations of odor quality between binary mixtures and their components*. Chem. Senses, 25, 247–265.

Woodworth, R.S. and **Schlosberg, H.** (1954) Experimental Psychology, revised edn. Holt, Rinehart & Winston, New York.

Woskow, M.H. (1968) *Multidimensional scaling of odors*. In Tanyolaç, N. (ed.), Theories of Odor and Odor Measurement. Robert College, Istanbul, pp. 147–191.

Wright, R.H. and **Michels, K.M.** (1964) *Evaluation of infrared relations to odor by a standards similarity method*. Ann. NY Acad. Sci., 116, 535–551.

Wysocki, C.J., Pierce, J.D. and **Gilbert, A.N.** (1991) *Geographic, cross-cultural, and individual variation in human olfaction*. In Getchell, T.V., Doty, R.L., Bartoshuk, L.M. and Snow, J.B. (eds), Smell and Taste in Health and Disease. Raven Press, New York, pp. 287–314.

Yokoi, M., Mori, K. and **Shigetada, N.** (1995) *Refinement of odor molecule tuning by dendrodendritic synaptic inhibition in the olfactory bulb*. Proc. Natl Acad. Sci. USA, 92, 3371–3375.

Yoshida, M. (1964a) *Studies in psychometric classification of odors (4).* Jap. Psychol. Res., 6, 115–124.

Yoshida, M. (1964b) *Studies in psychometric classification of odors (5)*. Jap. Psychol. Res., 6, 145–154.

Yoshida, M. (1972) *Studies in psychometric classification of odors (6)*. Jap. Psychol. Res., 14,70–86.

Yoshida, M. (1975) *Psychometric classification of odors*. Chem. Senses Flav., 1, 443–464.

Zwaardemaker, H. (1895) Die Physiologie de Geruchs. Engelmann, Leipzig.

Zwaardemaker, H. (1925) L'Odorat. Doin, Paris.

Accepted February 3, 2000