C(S)

Olfactory Quality: From Descriptor Profiles to Similarities

Pascal Callegari, Jacques Rouault and Paul Laffort

Groupe Physiologie de la Chimiorécéption, CNRS–IESGI–UPR 9054, ENSBANA, Campus Universitaire, 1 Esplanade Erasme, F-21000 Dijon, France

Correspondence to be sent to: Pascal Callegari, Groupe Physiologie de la Chimiorécéption, CNRS–IESGI–UPR 9054, ENSBANA, Campus Universitaire, 1 Esplanade Erasme, F-21000 Dijon, France

Abstract

The model described here, which links olfactory semantic profiles to canonical similarities, performs significantly better than the current state of the art, mainly due to the use of an optimization procedure. Application of Tversky's model in this study shows that in olfaction, estimation of similarities is based exclusively on the distinctive elements of the profiles, and does not take into account the common elements. Moreover, the optimum number of descriptors to reckon with seems to lie between 25 and 30. **Chem. Senses 22: 1–8, 1997.**

Introduction

The problem of describing objectively an odor perceived by humans has been addressed for a considerable time. Several methods have been used to address this problem.

(i) The use of profiles of reference odors: an odorous substance is described in terms of a similarity profile that is related to a certain number of reference substances considered to represent as truly as possible the olfactory space (this poses the problem of the choice of the reference substances; in other words, how can the validity of the choice be assured?). Several authors have used this technique, e.g. Wright and Michels (1964), Amoore and Venström (1965), Boelens and Haring (1980) and more recently Jaubert *et al.* (1987) as well as Takagi (1987).

(ii) The use of profiles of semantic descriptors: an odorous substance is described by means of a list of semantic descriptors, the values of which can vary in intensity. For instance, in Dravnieks *et al.* (1978) and Dravnieks (1985), using a list of 146 descriptors, the intensity varies from 0 (descriptor absent) to 5 (descriptor strongly present). It is particularly interesting to note that the results produced by this method are reproducible and especially stable (Dravnieks, 1982). This method has been adopted by the American Committee for normalization (ASTM E-1804) (Dravnieks, 1985).

(iii) The use of odor similarities: the likeness between two substances is ranked on a numerical scale fixed *a priori*. Many studies are based on this technique, e.g. Engen (1962), Schutz (1964), Yoshida (1964), Woskow (1968) Berglund *et al.* (1972), Dravnieks (1974) and Döving and Lange (1978).

We are particularly concerned here with the last two methods. Indeed, if they represent two aspects of the same reality, it is not unrealistic to think that one could conceive a mathematical model that allows one type of data to be transformed to the other.

A first attempt in this direction was made by Dravnieks *et al.* (1978), who used the Euclidian distance and the

chi-squared distance to predict indexes of dissimilarity starting from semantic profiles. (These two measures are

$$\text{Euclidian distance} = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2$$

and

$$\text{Chi-square } \chi^2 = \sum_{i=1}^{n}\left(\frac{x_i - y_i}{y_i}\right)^2$$

in which n is the size of the vector X, x_i the ith value of the vector X). The correlations obtained are of the order of 0.8 for a sample of 10 substances, which is too small a number to draw general conclusions from.

Rouault and Laffort (1993) proposed the use of 11 families of distances or proximities to transform a semantic description (Dravnieks, 1985) into an odor similarity. The principal conclusion concerned the value of the exponent of these formulas, which took on values between 0.1 and 5.0; these authors observed an optimization of the correlations between calculated values of similarity and those found measured in the literature for an exponent close to 1.0. On the other hand, uncertainty remained as to the best performing family of distances, as some of them appeared equivalent (the correlations varied from 0.73 to 0.78).

The construction of a model to switch between semantic profiles and odorant similarities not only meets a real need to obtain rapidly a great number of similarities starting from a restricted number of profiles, but can also give us effective information about the nature of the cognitive processes that allow us to establish similarities between objects in the outside world. Indeed, schematically one could consider that the brain works in two different modes: analytical and synthetic. The analytical mode is expressed by a precise description of a semantic character (each character being individually and quantitavely evaluated). The synthetic mode expresses itself by a global response graduated on a predefined scale that can even be binary (either there is a resemblence, or there is not). The interest in determining a mathematical model that enables one to go from the analytical to the synthetic mode is evident immediately. In fact, if one wishes to compare two by two a great number n of objects, in the synthetic mode one has to proceed to $n(n-1)/2$ comparisons, while in the analytical mode only np descriptions are necessary, where p is the number of descriptors.

The first step consisted of defining a strongly coherent group of experimental data on similarities, on which to construct the predictive model.

Determination of a coherent group of authors

Callegari and Rouault (1994) have shown that similarity and dissimilarity are not equivalent (or symmetrical) on the basis of repetition data in the literature, and that it is preferable to work with dissimilarity data, while the similarities can be made homogeneous with dissimilarities through a transformation of the square root type. They defined this type of homogeneous data using the term 'canonical similarity'.

First, a group of authors was established whose values of canonical similarity are coherent. One begins by comparing two by two the different publications (the similarities and dissimilarities are normalized and homogenized to one distance as indicated above). For each pair of publications providing at least three common canonical similarities (i.e. corresponding to the same pairs of substances), we calculate both the Pearson and Spearman correlation coefficients. The results are given in Table 1. The authors used in this part are: Amoore and Venström (1965) [AV], Berglund *et al.* (1973) [BB], Boelens and Haring (1980) [BO], Döving and Lange (1966) [DL], Dravnieks *et al.* (1978) [D8], Dravnieks (1974) [D4], Engen (1962) [EN], Schutz (1964) [SC], Woskow (1968) [WO], Wright and Michels (1964) [WM], Yoshida (1964, table 3) [Y3], Yoshida (1964, table 7) [Y7].

Next, the significance of these correlations is evaluated with the help of the MinPS statistic (Rouault and Laffort, 1993): the degree of significance is the smallest significance produced by each of the two coefficients of correlation. The joint use of these two coefficients of correlation and this logic of the smallest value allows us to ascertain the coherence of the significance of the results. These are reported in Table 2.

To select the group of coherent authors, the following criteria were applied. (i) All authors never connected by a strong significance ($P < 0.01$) to another author (Table 2) were eliminated. This removed AV, DL and D4. (ii) Authors whose strong significances correspond to at least one of the values of Pearson or Spearman below 0.8 (Table 1) were eliminated; BB, BO, EN, SC, Y3 and Y7 were removed by this criterion.

Table 1 Between-author correlations of canonical similarities

	AV	BB	BO	DL	D4	D8	EN	SC	WM	WO	Y3	Y7
AV	107 749											
BB	9 0	21 210										
BO	35 16 0.408 0.450	11 40 0.266 0.450	342 9360									
DL	7 0	2 1	5 2	10 18								
D4	8 0	2 1	2 0	1 0	20 190							
D8	4 0	4 24 0.626 0.602	5 20 0.785 0.694	1 0	1 0	10 220						
EN	1 0	2 1	1 0	0 0	1 0	3 12 0.594 0.500	5 15					
SC	7 0	9 40 0.691 0.693	9 28 0.467 0.472	1 0	3 4 0.971 0.900	3 16 −0.031 0.035	4 6 0.993 0.957	19 180				
WM	13 0	9 20 0.473 0.728	21 37 0.535 0.470	3 3 0.974 0.875	8 0	2 0	2 0	8 0	51 450			
WO	8 0	10 45 0.478 0.426	15 56 0.352 0.443	4 1	4 6 0.533 0.486	5 40 0.907 0.908	6 15 0.778 0.805	11 48 0.705 0.557	16 48 0.818 0.813	25 300		
Y3	6 0	1 0	13 66 0.370 0.463	2 0	0 0	0 0	1 0	2 2	8 0	6 15 0.477 0.360	20 190	
Y7	3 0	3 3 0.500 0.625	19 103 0.319 0.327	3 1	0 0	0 0	0 0	1 0	7 12 0.614 0.406	5 10 0.748 0.691	4 6 0.553 0.429	24 276

Each box, off the diagonal, contains on the top line at the left side the number of common substances, at the right side the number of common similarities, and on the second and third line, the Pearson and the Spearman correlation coefficients respectively. The diagonal contains from the top downwards for each author the number of substances and the number of similarities.

In the end, one thus gets a homogeneous group limited to three authors, D8, WM and WO, referred to as M3. This procedure is probably too severe, but a strong coherence of basic data is necessary in order to optimize satisfactorily a model that will enable us to switch from semantic profile to canonical similarity.

From profiles to similarities

The goal of this step is twofold. In the first instance, it is meant to propose a reliable model that enables canonical similarities to be predicted from semantic profiles of odor description. Secondly, the aim is to determine the minimal

Table 2 Significance table of the correlations of Table 1 following the minimal logic MinPS explained in the text

	AV	BB	BO	DL	D4	D8	EN	SC	WM	WO	Y3
BB											
BO	NS	NS									
DL											
D4											
D8		0.01	0.01								
EN						0.10					
SC		0.01	0.02		NS	NS	0.02				
WM		0.05	0.01	NS							
WO		0.01	0.01		NS	0.01	0.01	0.01	0.01		
Y3			0.01							NS	
Y7		NS	0.01						NS	0.05	NS

Table 3 Evolution of the correlation between observed values and values predicted by multiple linear regression, in connection with the threshold value (expressed in %)

Threshold value	0	5	10	15	20	25	30	35	40
Correlation	0.567	0.795	0.781	0.739	0.745	0.756	0.751	0.679	0.577

number of descriptors that allows an odor to be described as completely as possible.

A variant of the model of Tversky

Tversky (1977) proposed a qualitative model that defines the similarity between two objects A and B:

$$S(A,B) = \alpha f(A \cap B) + \beta f(A - B) + \delta f(B - A) \quad (\text{with } \alpha > 0,\ \beta < 0,\ \delta < 0)$$

The similarity S between two objects A and B is made equal to a linear combination of the traits common to A and B, the traits specific to A and those specific to B, expressed as $f(A \cap B)$, $f(A - B)$ and $f(B - A)$ respectively. The appeal of this model lies in its intuitive acceptability.

A first modification to this basic model is to put $\beta = \delta$, for nothing justifies the assumption that a distinctive character proper to A is more informative than a distinctive character proper to B. In the event of a decision to be made on this, one should have the original similarity data and, above all, know the order of presentation of the odors.

The model thus becomes:

$$S(A,B) = \alpha f(A \cap B) + \beta[f(A - B) + f(B - A)] \quad (\text{with } \alpha > 0,\ \beta < 0)$$

In the present case, the function f has been chosen as a ranking function that counts the number of descriptors present. For example, one counts the number of descriptors present at the same time both for A and B.

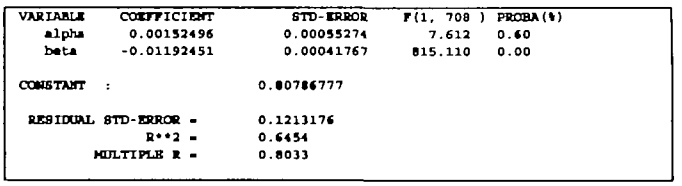

Figure 1 Result of the multiple linear regression with the threshold fixed at 4%.

The problem now is to establish a threshold value below which the descriptor is considered to be absent. This amounts to binarizing the profile values earlier established between 0 and 100. Next, a multiple linear regression is used between the values deduced from the binarized profiles and the canonical similarities of M3 in order to determine the optimal value of the threshold. The results are presented in Table 3 and Figure 1.

It seems in Table 3 that the optimum lies in the neighbourhood of 5%. A finer analysis yields a more precise value of 4% (Figure 1). This value is remarkable for it corresponds exactly with the threshold of significance for the percentage applicability established by Dravnieks (1985). [It should be specified here that each descriptor is affected

by a percentage of applicability that corresponds to a balance between the number of subjects using this descriptor and the perceived intensity scaled from 0 to 5. The percentage of applicability takes on values between 0 and 100% according to Dravnieks (1985).]

One can draw another interesting conclusion from Figure 1, when observing the values of the regression coefficients α and β; indeed, one can see the preponderance of β (ratio 8/1), confirmed by the value obtained by the *F*-test (ratio 107/1). This shows the dominance of the distinctive characters over the common characters. This result can be confirmed by not taking into account the common characters: one ends up with the same correlation value. This would tend to prove that in the first approximation one could ignore the common traits and retain only the distinctive traits.

It can be noted that the model derived from Tversky's work as it emerges from the linear regression is not completely satisfactory for two reasons. In the first place, the model is not valid at the borders (when all elements are common or distinctive, the obtained values are not respectively 0 and 1); secondly, the linear regression produces a constant (see Figure 1) to which no meaning can be given.

Therefore, a new model that does not have these disadvantages is proposed:

$$S(A,B) = 1 - \frac{\text{Distinctive}(A,B)}{n}$$

Distinctive(A,B) is the count of the number of descriptors that are distinct for A and B among the *n* descriptors of the profile, i.e.

$$\text{Distinctive}(A,B) = f(A-B) + f(B-A)$$

This formula is valid at the borders. If indeed all *n* descriptors are distinct, the similarity has a value of 0; if all are common, or by the same token, all absent, *S* will take on the maximum value of 1. It should be observed that this method is analogous to the use of a city-block distance

$$\text{CTB} = \sum_i |X_i - Y_i|$$

when binary values 0 and 1 are used.

Reduction of the global number of descriptors

The method used can be summarized as follows. Of the 146 descriptors, the first is withdrawn, the similarities derived without this descriptor are calculated, the correlation between derived and experimental similarities is calculated. Next, the second, the third, etc., are treated in the same way. Then, the descriptor giving the strongest correlation upon its withdrawal is suppressed. This process is repeated for the suppression of all descriptors.

This algorithm is a method of stepwise optimization. The results are given in Figure 2 and show the evolution of the correlation between predicted and observed values as a function of the number of descriptors taken into account. It can be observed that reducing the number of descriptors from 146 to 101 (zone I) lets the correlation emerge from the noise, and corresponds to the redundancy of certain descriptors (the correlation grows from 0.81 to 0.87); from 100 to 26 descriptors (zone II), it remains constant (equal to 0.87); finally, from 25 to 1 descriptors (zone III) the correlation falls from 0.87 to 0.56. This fairly clearly shows that 25 well-chosen descriptors seem sufficient to represent faithfully the perceptual olfactory space.

Discussion

The main importance of this study (which was not initially suspected) is to show that in olfaction, contrary to what Tversky's model predicts, the estimation of a similarity is based only on the distinctive elements; the common elements do not add any information. *A posteriori*, this result is not very surprising, to the extent that the sensory systems, generally, are programmed to estimate contrasts, rather than absolute values. What is true (and well known) in the constitution of a visual or olfactive image, remains true according to the present work in the comparison of two olfactive images. The question of the possible generalization of this fact to other sensory modalities is thus open.

To this first result is added the one initially pursued, namely an important gain in correlation, when compared with earlier work, between values of experimental olfactory similarities and values of similarities derived from semantic profiles (for an approximately equal number of data, i.e. 711 similarities, the Pearson correlation coefficient increases from 0.78 to 0.87). The question remained as to which part of the gain was due to the optimization procedure, and which part to the switch model, properly called the 'modified Tversky model'. With this in mind, four of the families of mathematical distances that were found to be the

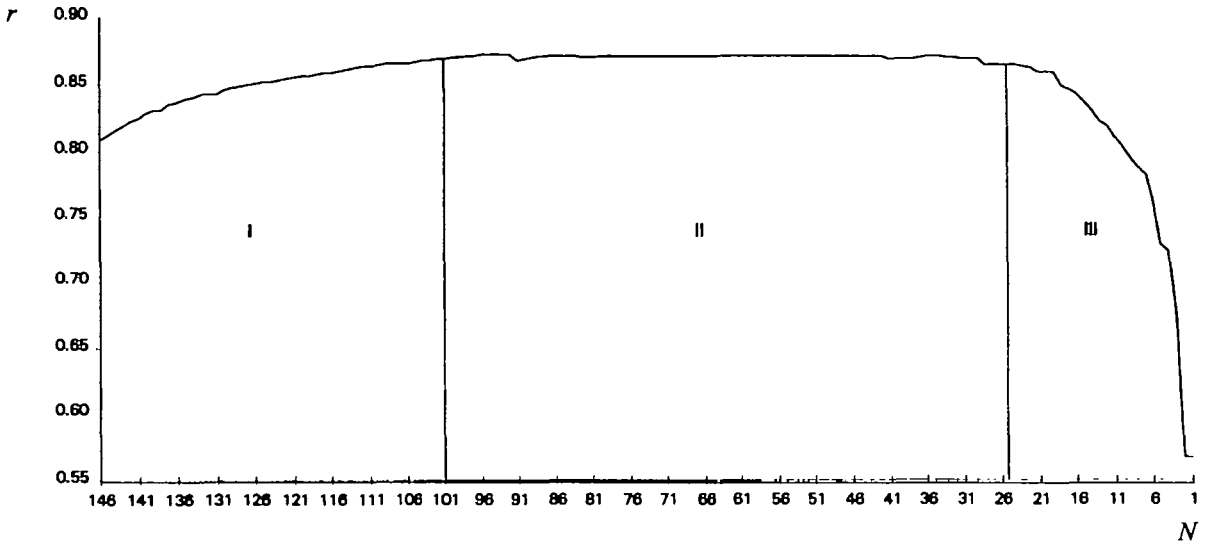

Figure 2 Evolution of the correlation *r* between canonical experimental similarities and similarities calculated following the 'modified Tversky model', as to the number of descriptors *N* (see text).

best performing by Rouault and Laffort (1993) have been submitted to the same optimization procedure as described above, for different values of the exponent *n*. The distances studied analyzed were those of Rouault (ROU), Minkowski (MIN), Benzecri (BEN) and Tanimoto option 1 (TA1) defined respectively by the relations:

$$\mathrm{ROU}_n = \frac{1}{n} \sqrt[n]{\frac{1}{N} \sum_i \left| \frac{x_i}{\bar{x}} - \frac{y_i}{\bar{y}} \right|^n}$$

$$\mathrm{MIN}_n = \sqrt[n]{\sum_i |x_i - y_i|^n}$$

$$\mathrm{BEN}_n = \sqrt[n]{N \sum_i \frac{\sum_i x_i + \sum_i y_i}{x_i + y_i} \times \left| \frac{x_i}{\sum_i x_i} - \frac{y_i}{\sum_i y_i} \right|^n}$$

$$\mathrm{TA1}_n = 1 - \sqrt[n]{\frac{\sum_i [\inf(x_i, y_i)^n]}{\sum_i [\sup(x_i, y_i)^n]}}$$

in which $\bar{x}$ represents the mean value of the profile X, N the size of the vector X, x_i the ith value of the vector X, and n is the exponent.

The results are obtained in the form of three-dimensional curves showing the variation of the correlation as a function

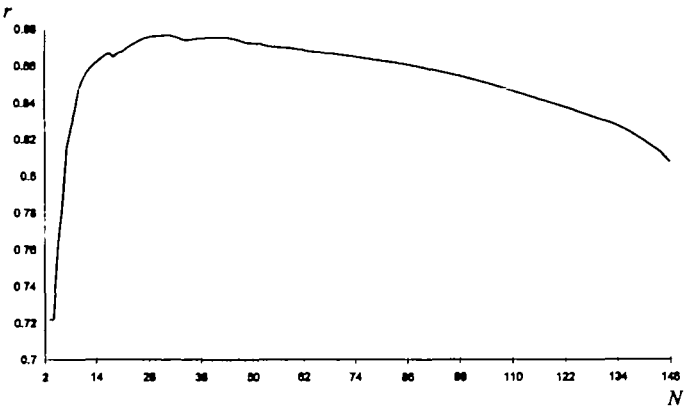

Figure 3 Representation of the variation in Pearson correlation coefficient *r* obtained with the model ROU (exponent = 0.7), as a function of the number of semantic descriptors *N*, between experimental similarities and those derived from semantic profiles (number of similarity data: 711). The maximum of Pearson correlation coefficient is obtained for 30 descriptors.

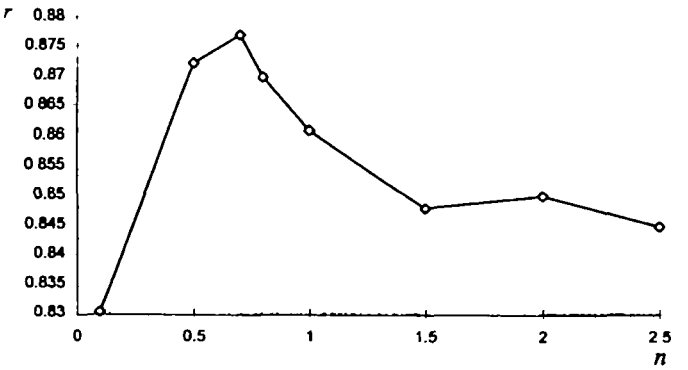

Figure 4 Representation of the variation in Pearson correlation coefficient *r* obtained with the model ROU, as a function of the values of the exponent *n* and for 30 semantic descriptors, between experimental similarities and those derived from semantic profiles (number of similarity data: 711).

of n and the number of descriptors. Orthogonal sections of this kind of three-dimensional figures, in the particular case of the model ROU, are reproduced in Figures 3 and 4. They allow us to verify that, for this family of distances, the variation of the exponent is accompanied by virtually constant values of the results, in terms of the value of the correlation coefficient, for values of n between 1.5 and 2.5 (r = 0.84–0.85), and presents an optimum between 0.5 and 0.8 (r = 0.87–0.88). On the other hand, the optimal correlation value is obtained when the number of descriptors is 30, the curve descending regularly from 30 to 146, without a plateau as observed with the modified Tversky model. The other models show analogous courses, demonstrating that the gain in correlation is principally due to the optimization procedure, with no possibility of a well-considered choice between the different models of distance (the maximum of r varies between 0.85 and 0.88, the last value coming from the model ROU).

Concerning the number of relevant, well-chosen descriptors, the number of 25–30 can be retained, which is indeed about midway between the values obtained by different authors: 17 by Jeltema and Southwick (1987), 30 by M. Chastrette (personal communication), 40 by Jaubert and Doré (1987).

The application of this approach to the profiles of reference odors, mentioned in the Introduction (the third way to evaluate a quality), will be the object of a later study.

ACKNOWLEDGEMENTS

The authors thank sincerely Hans van der Starre for his help in editing the English text. This study benefited from financial support from the Régie Autonome des Transports Parisiens (RATP) (project studying odor on the Paris metro) and from the Conseil Régional de Bourgogne.

REFERENCES

Amoore, J.E., Venström D. (1965) Correlations between stereo-chemical assessments and organoleptic analysis of odorous compounds. In Hayashi, T. (ed.), *Olfaction and Taste 2*. Macmillan, New York, pp. 3–17.

Berglund, B., Berglund, U., Engen, T. and Ekman, G. (1972) Multidimensional analysis of twenty-one odors. Reports from the Psychological Laboratories, University of Stockolm, no. 346.

Boelens, H. and Haring, H.G. (1980) *Molecular Structure and Olfactive Quality*. Narden, Bussum, The Netherlands.

Callegari, P. and Rouault J. (1995) Are odour similarities and dissimilarities experimentally symmetrical ? ECRO-XI Abstracts, *Chem. Senses*, **20**, 99.

Döving, K.B. and Lange, A. (1966) Comparative studies of sensory relatedness of odours. Reports from the Psychological Laboratories, University of Stockolm, no. 211.

Dravnieks, A. (1974) A building-block model for the characterization of odorant molecules and their odors. In Cain, W.S. (ed.), *Odors: Evaluation, Utilisation and Control. Ann. NY Acad. Sci.*, **237**, 144–163.

Dravnieks, A. (1982) Odor quality: aemantically generated multidimensional profiles are stable. *Science*, **218**, 799–801.

Dravnieks A. (1985) *Atlas of Odor Character Profiles*. ASTM Data Series 61, Philadelphia.

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

Engen T. (1962) The psychophysical similarity of the odors of aliphatic alcohols. Reports from the Psychological Laboratories, University of Stockholm, no. 127.

Jaubert, J.-N., Gordon, G. and Doré, J.-C. (1987) Une organisation du champ des odeurs. *Parf. Cosm. Arômes*, **77**, 53–56.

Jeltema, M. and Southwick, R. (1986) Evaluation and application of odor profiling. *J. Sens. Stud.*, **1**, 123–136.

Rouault, J. and Laffort, P. (1993) Le cerveau olfactif utiliserait-il une distance mathématique? In Boussaid, O., Brissaud, M., Ritschard, G. and Royet, J.-P. (eds), *Pluridisciplinarité dans les sciences cognitives*. Hermès, Paris, pp. 296–286.

Schutz, H.G. (1964) A matching-standard method for characterizing odor qualities. In Whipple, H.E. (ed.), *Recent Advances in Odor: Theory, Measurement and Control. Ann. NY Acad. Sci.*, **116**, 517–526.

Takagi, S.F. (1987) A standardized olfactometer in Japan. A review

over ten years. *In* Roper, S.D. and Atema, J. (eds), *Olfaction and Taste IX. Ann. NY Acad. Sci.*, **510**, 113–118.

Tversky, A. (1977) Features of similarity. *Psychol. Rev.*, **84**, 327–352.

Woskow, M.H. (1968) Multidimensional scaling of odors. *In* Tanyolaç N.N. (ed.), *Theories Of Odor and Odor Measurement*. pp. 147–188.

Wright, R.H. and Michels, K.M. (1964) Evaluation of infrared relations to odor by a standards similarity method. In Whipple, H.E. (ed.), *Recent Advances in Odor: Theory, Measurement and Control*, *Ann. NY Acad. Sci.*, **116**, 535–551.

Yoshida, M. (1964) Studies in psychometric classification of odors. *Jpn. J. Psychol.*, **35**, 1–17 (in Japanese).

Received on November 21, 1995; accepted on July 25, 1996