Cs

Is Short-term Odour Recognition Predictable from Odour Profile?

J.-P. Royet, H. Paugam-Moisy[1], C. Rouby, D. Zighed[2], N. Nicoloyannis[3], S. Amghar[3] and G. Sicard

Laboratoire Neuroscience et Olfaction, Université Claude-Bernard Lyon 1 (CNRS ERS 5643), 69622 Villeurbanne, [1]Laboratoire de l'Informatique du Parallélisme, Ecole Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, [2]Equipe de Recherche en Ingénierie de la Connaissance, Université Lumière Lyon 2, 69500 Bron and [3]Laboratoire des Méthodes d'Analyse des Systèmes et des Structures (URA 934 CNRS) Université Claude-Bernard Lyon 1, 69622 Villeurbanne Cedex, France

Correspondence to be sent to J.P. Royet, Laboratoire Neuroscience et Olfaction, Université Claude-Bernard (ERS 5643 CNRS), 69622 Villeurbanne, France

Abstract

Smelling an odour induces a pattern of sensations, images and memories which participate in identification. It was proposed that perceptual memory performances for odours could be inferred from the description of these olfactory representations. The subject was asked to elaborate an odour descriptor profile, and a short-term odour recognition memory task was chosen to test the individual perceptual memory performance. Two pattern-recognition methods based on artificial neural networks and discriminant analysis were carried out and allowed odour profile and perceptual memory performance to be related. Insofar as the subjects gave dichotomic responses in the recognition memory task, each response could be evaluated in terms of correct or incorrect responses. Simulations indicated that the olfactory recognition memory performance can be predicted in man from odour-elicited semantic profiles by using artificial neural networks. It was also shown that all semantic descriptors do not participate in olfactory recognition to the same degree. Low-level information, such as intensity, familiarity and hedonic judgements, did not allow the artificial neural network to predict the olfactory performance. By contrast, high-level information, such as gustatory, olfactory and visual evocations, allowed artificial networks to make such predictions. **Chem. Senses 21: 553–566, 1996.**

Introduction

It has long been shown that most people have difficulty in naming odours (Lawless and Engen, 1977). They are often able to recognize a smell as being familiar without being able to assign it a correct label. Lawless and Engen (1977) named this state 'tip-of-the-nose' phenomenon, by analogy to the 'tip-of-the-tongue' phenomenon (Brown and McNeill, 1966). These authors clearly showed that although subjects could not name the odour, they could name a similar odour, an object evoking the target odour, a general category for the smell, a place the smell might have come from, and formed a visual image of the object or the place. Thus, they could provide a pattern of semantic descriptors from their personal sensory evocations. Schab (1991) considered that olfactory identification varies in informational specificity from the most primitive levels of judgement (i.e. pleasantness, intensity, familiarity) to single-label, object-name identification with various intermediate steps. These could be generic judgements (e.g. a floral note that does not allow precise identification of the corresponding flower), evocations of a non-olfactory sensation (e.g. acid, sweet, heavy, etc.) or memories of a place (e.g. hospital, kitchen, school, etc.). Thus, unlike sounds and images, odours are generally not stored in memory as unique entities, but are frequently associated with other sensory perceptions such as gustatory, tactile, visual, auditory and thermal sensations.

Odour profiling is an attempt to characterize odorous stimuli by profiles of their individual qualities, components or 'notes' (Doty, 1991). This procedure has often been employed in the description of food aromas (Moskowitz and Barbe, 1976; Moskowitz and Gerbers, 1974). Dravnieks (1982) showed that profiles based on the responses of a large number (e.g. 150) of panellists are stable representations of the odour character and are robust constructs. Another procedure currently used by experimenters to characterize odorants is proximity analysis, which reduces linguistic influence in odour classification. The proximity analyses are based upon the rating of dissimilarity between odorants from multidimensional scaling. According to Moskowitz and Gerbers (1974), their advantage is that the observer is allowed to generate his or her own dimensions. Descriptions of odorants based on odour profiling and measurements of qualitative dissimilarity from proximity analyses are two aspects of the subjective judgements of odours. It is possible to predict the second one from the first one by performing multiple linear regressions in salience analyses (Moskowitz, 1974). However, the proximity analysis approach to odour quality does not allow validation of the correctness of the subject's responses. Other measurements, based on discrimination or short-term recognition memory, in which the odours are presented by pairs, can be used to take the quality of the judgements into account objectively. As the subjects indicate that the two odours are identical or different, the judgements can be evaluated in terms of correct or incorrect responses.

We hypothesized that a relationship may exist between the semantic aspects of the odour description and the olfactory recognition performance: the individuals who describe odours correctly also recognize these odours correctly. This hypothesis has several corollaries. First, if we can evaluate the accuracy of odour description, we can then rate the correctness of recognition performance. Second, a correct description requires a sufficient amount of semantic information; if not, odour recognition is not possible. Third, the semantic data may have not equal informational meanings. For example, intensity and familiarity can be considered to be less important in odour recognition than veridical label (Schab, 1991).

The aim of the present experiment was to elaborate a heuristic method that would allow one to determine profiles representative of odours and to predict the subjects' recognition performances from their patterns of semantic descriptors. In order to evaluate the olfactory performance, we used a very short-term odour recognition task presenting paired odours with a retention interval of 30 s. The subject's performances in these conditions have been previously well characterized (Jehl *et al.*, 1994, 1995). In order to evaluate the descriptive quality, we used the odour profiling method. Finally, in order to predict the olfactory performance from odour descriptions, we used two kinds of methods: artificial neural networks and discriminant analyses. In contrast to multiple linear regression analysis, these methods can be applied to qualitative values.

Materials and methods

Subjects

One hundred and five naive subjects, 47 men (mean age 34.74 ± 9.96 years) and 58 women (mean age 35.98 ± 12.18 years) participated to this experiment. They were all French speaking, although 14 subjects (13.3%) were from French-speaking Arabic countries (Morocco, Tunisia and Algeria). None presented olfactory symptoms such as hyposmia or anosmia. The subjects were asked not to use perfumes and not to eat chewing gum before testing. They were paid for their participation.

Stimuli

The stimulus sample included 17 odorants (Table 1). These were pure chemicals and were selected so as to get very or slightly different odour qualities. Odours were contained in 250 ml yellow glass jars with polypropylene screw lids (OSI, France). Jars were opaque to mask any visual identification cues. Each bottle contained 10 ml of a 1% solution, except for 2,3-dimethyl-2,3-butanediol for which a 10% solution was used. Diethyl phthalate served as solvent for dilution.

Table 1 Order of presentation of nine odour pairs with corresponding similarity level

Pair	Odour	Odour	Similarity
0[a]	anisole	anethole	very different
1	eugenol	limonene	very different
2	hexenol	hexenol	identical
3	coumarin	coumarin	identical
4	geraniol	phenylethyl alcohol	very different
5	naphthalene	naphthalene	identical
6	aldehyde C8	aldehyde C10	slightly different
7	3,3-dimethyl-2-butanone	2,3-dimethyl-2,3-butanediol	slightly different
8	menthone	*l*-carvone	very different
9	β-hexenyl acetate	iso-amyl acetate	slightly different

[a]Odour pair used to familiarize subjects with the experimental procedure.

Olfactory descriptors

In order to obtain an odour quality estimation from the subjects, we examined the relevance of 140 descriptors distributed in eight categories. The first two categories (each consisting in a single descriptor) yielded the degree of familiarity and intensity of each odour in a pair and were rated as a bipolar continuum with a scale from 0 to 9. Categories 3–7 were composed of 132 French semantic descriptors from which the subject had to select the most pertinent descriptors (Table 2) of the second odour of each pair. Multiple-choice tests reduced the tip-of-the-nose phenomenon to a minimum. These five categories consisted of hedonic terms, olfactory terms, evocations of non-olfactory sensations (gustatory and somesthesic), the veridical labels, and memories of places. In each category, the descriptors were given in alphabetical order. In addition, for each semantic category, the subject could give another personal description as indicated in the list by terms 'other-hedonic', 'other-olfactory', 'other-nonolfactory', 'other-veridical' and 'other-place'. Thus, 132 French semantic descriptors from a list and five personal descriptions, one per category, were available. Finally, the subject could describe his or her personal memories (category 8) through anecdotes. The souvenir had to be specified in terms of time, place and possibly related people. The anecdote of the 'madeleine' of Marcel Proust was described as an example at the beginning of the session.

Experimental procedure

Nine different odour pairs were successively presented (Table 1). Three pairs were made with identical odours, three pairs were made with very different odours, and three pairs were made with slightly different odours. The criterion of proximity between odours was chosen in terms of ease of discrimination. To measure odour profiling, we were interested in the absolute intrinsic identification performances of the subjects, not in the relative performance of each odorant. Therefore, only the description of one of the odours in a pair seemed to be useful. In addition, so as not to interfere with the odour recognition memory task, only the second odour was described with odour profiling. Rather then letting the subjects remember the labels they attached to the odours before testing recognition performances, we would rather have them remember odours only from a perceptual point of view. Obviously, this kind of procedure does not prevent the subjects from spontaneously performing some cognitive processing for the first odour (Schab, 1991). However, this uncontrolled bias was all the more limited because the interval between both odours was short and because the subjects also performed a low-level processing during this interval (judgements of intensity and familiarity).

The experimenter presented the first odour of the first pair to the subject with an instruction to sniff for 5 s. The subject was asked not to spill the glass jars and to sniff without letting the mouth of the jars touch any part of his/her body such as the nose or the lips. The subject determined the degree of intensity and familiarity of the odour. Thirty seconds after presentation of the first odour, the second odorant of the pair was presented to the subject, who had to decide whether both odorants were identical or not. Then, he/she had to rate intensity and familiarity, and

Table 2 List of semantic descriptors (translated from French) used for categories 3–7

	Hedonic	Olfactory	Gustatory and somesthesic	Veridical labels		Place
1	pleasant	alcoholic	sour	garlic	gas	cupboard
2	appetizing	ammoniated	pungent	almond	tar	washbasin
3	unpleasant	animal	acid	anaesthetic	grass	cafeteria
4	nauseating	woody	bitter	aniseed	smoked ham	cellar
5	neutral	burning	tart	mothkiller	lavender	field
6	stinking	camphoraceous	artificial	citronella	vegetable	butcher's shop
7	repugnant	spicy	heady1	banana	cleaning liquid	kitchen
8	other-hedonic	ethereal	warmly	butter	medication	dentist
9		faecal	chemical	wood	mint	school
10		flowery	faded	firewood	motor	sewer
11		fruity	sweet	candy	hazelnut	cowshed
12		smoked	heady2	candle	onion	garage mechanic
13		herbaceous	insipid	cadaver	cosmetic	barn
14		human	fresh	cinnamon	painting	hospital
15		medicinal	cold	cherry	pipe	maternal school
16		putrid	oily	mushroom	smoked fish	house
17		sulphurous	irritating	chewing-gum	pepper	sea
18		earthy	light	lemon	apple	lawn
19		plant	heavy	clove	liquorice	underwood
20		other-olfactory	metallic	paste	rose	dyeing
21			natural	leather	soap	other-place
22			piquant	disinfectant	sulphur	
23			rancid	eau de cologne	sweat	
24			dry	incense	tea	
25			sugared	turpentine	urine	
26			other-nonolfactory	eucalyptus	vanilla	
27				ether	varnish	
28				hay	freshmeat	
29				cheese	vinegar	
30				manure	violet	
31				tart	other-veridical	

Heady1, heady2 are discriminative terms that could not be found in English.

to select the most pertinent descriptors of the second odour. An individual response sheet was given to the subjects to write down their recognition judgement (identical or different), the different descriptors selected in each category and the memory. The number of descriptors was limited to two terms when the list was short (hedonic, olfactory and place) or three terms when the list was long (non-olfactory or veridical). Only the running number of the descriptors in the list was indicated in the response sheet. Four minutes were needed to record the responses of the subjects before another pair was presented.

The choice of a retention interval of 30 s was a compromise between the necessity to observe a reduction in the olfactory performance, to record the semantic responses from the subjects, but also to limit the duration of the experiment. It has been observed that very short-term recognition memory performance was significantly reduced to 30 s as a function of the qualitative distance between odours (Jehl *et al.*, 1994), but also to 20 s as a function of familiarization (Jehl *et al.*, 1995).

To familiarize the subjects with the experimental conditions, an initial odour pair (trial pair) was presented at the beginning of the experiment. In addition, subjects were familiarized with the list of descriptors before testing began. They could ask the experimenter any question in order to

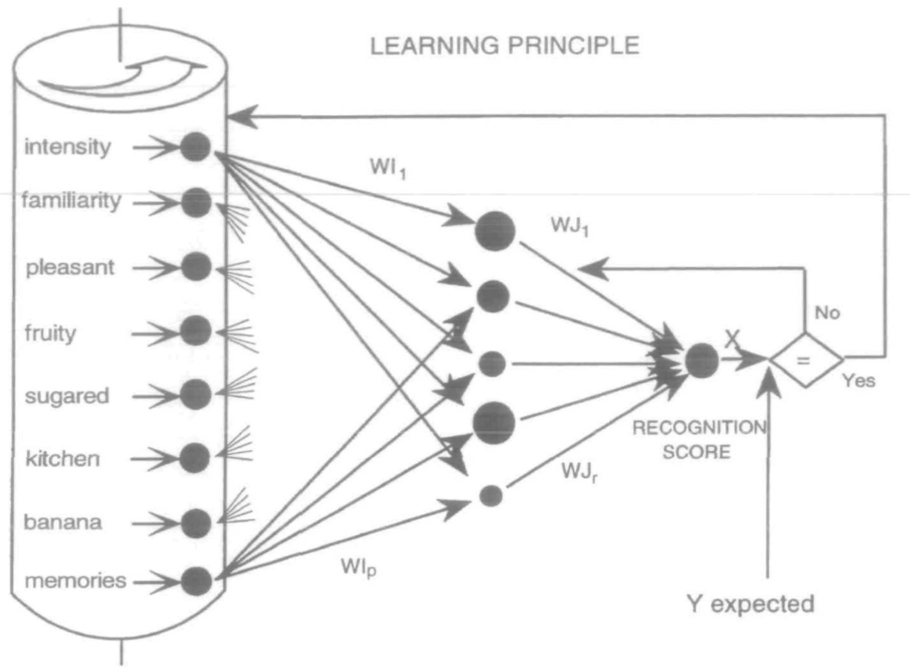

Figure 1 Learning principle of neural networks to mimic experimental data by associating recognition score with olfactory representations (intensity, familiarity, hedonic descriptors as pleasant, olfactory descriptors as fruity, non-olfactory descriptors as sugared, veridical descriptors as banana, descriptors of place as kitchen, and memories). For each odour and each individual, internal weights (W) of connections are modified by back-propagation (No) in order to reduce error between the calculated value (X), which is the network outputs, and the expected value (Y) which was experimentally obtained. When the threshold error is reached (Yes), evocation patterns of another individual are analysed.

explain any term. The order of presentation of the other pairs and the order of the two odours in a pair were randomly chosen at the beginning of the experiment, but were identical for all the subjects. The experimenter did not inform the subjects of the number of odour pairs to be tested. A complete test session, including presentation of all odour pairs, lasted for 1 h.

Modelling procedure

To model the process of odour recognition, we used connectionist methods of pattern recognition. A multilayer neural network is constituted of at least three layers of processing units: a layer of input units, a hidden layer consisting of 'feature detectors' and an output layer (Dayhoff, 1990). Values between processing units are weighted by a connection strength and the network is able to adjust them to associate input and output data. Two successive phases were performed: a learning phase and a generalization phase. For multilayer networks, the learning paradigm is the well-known error back-propagation rule (Rumelhart *et al.*, 1986). In the learning phase, the neural network is presented with a training set consisting of a set of pairs of patterns, each being an input pattern paired with a target output. For each pair of patterns, a single feedforward, a calculation of the quadratic error between the computed and the experimental outputs, and an adjustment of the values of the internal parameters (weights) by back-propagation are performed. When the whole set of pairs of patterns has been presented, the same procedure is applied again by presenting the whole set a second time. Several hundred presentations of the whole set can be required. Training is successful when the difference, or error, between the network's computed output and the target output is minimized. In the generalization phase, or test phase, the model can be validated by presenting new inputs that are different from the inputs of the training set. The performances of the neural network can be estimated by comparing the resulting computed outputs to the experimental target outputs.

In the present study, the neural network paradigm used was a fully interconnected back-propagation algorithm (Figure 1). Data from 84 subjects were involved in the learning phase, and those of 21 others were used for the generalization phase. The choice of the subjects for the generalization phase was performed according a jack-knife procedure (Tukey, 1958). The 105 subjects were subdivided into five groups of 21 subjects selected at random. Data from four groups were used for the learning phase and those

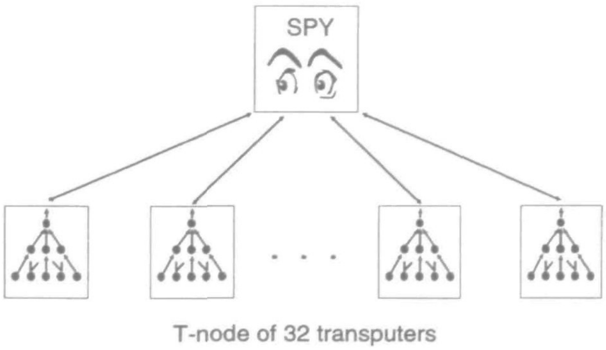

Figure 2 The spy-process and its farm of neural networks disposed according to a linear topology and implemented on a T-node of 32 transputers.

from one group were required for the generalization phase. The performances of the neural network were then computed for each odour pair. The successive use of data from four other groups for the generalization phase yielded four other sets of performance of the neural network. Mean performances were then computed for each odour pair. One multilayer network characterized by one input layer, one or several hidden layers and one output layer was dedicated to each odour pair. The input layer allowed us to take into account the occurrence of semantic descriptors. The output layer allowed us to represent the subjects' recognition scores. To perform the connectionist analyses, only the most frequently used descriptors (frequency ≥7) were retained. As a consequence, only the results from 27–35 descriptors per odour were analysed. The effective occurrence of a semantic descriptor and of a souvenir was scored 1; its absence was scored 0. The recognition scores were also coded 1 (correct response) or 0 (incorrect response).

To model the olfactory recognition performance optimally, the numbers of hidden layers and cells per layer had to be determined. In addition, testing each architecture of neural network can require several hundred iterations for learning. Therefore, the tuning of parameters required very high computational power. To explore and choose the best architecture and parameters of the network, a parallel computer was used. An optimization tool, known as 'a spy of parallel neural networks' (Figure 2) allowed the best performances to be surrounded automatically (Paugam-Moisy, 1991). In practise, the spy-process observed a set of neural networks, which all worked in parallel from the same set of patterns, but differed from one another in some of their parameters. The spy-process then computed an optimization algorithm for automatically selecting the neural networks which provided the best performance in terms of learning. This process was implemented on a T-node of 32 transputers. The two parameters to be opti- mized were the number of hidden units on one hidden layer, and a constant coefficient $K\alpha$:

$$K\alpha = \text{learning rate/fan-in}$$

where fan-in is the number of input links of each unit and the learning rate is the coefficient of weight updating in the back-propagation algorithm (LeCun, 1985; Rumelhart *et al.*, 1986). In order to provide results to be compared for the nine odour pairs, a network with only one hidden layer was considered in each case, although a few results were better with two hidden layer networks.

Finally, another method of pattern recognition was used and applied to the same data: discriminant analysis (Fisher, 1936), which is based on a class separation hypothesis which considers that individuals having distinct labels are represented in different areas separated by geometric borders of linear or ellipsoidal types. A learning phase means one can search, for the whole representation set, surfaces such as hyperplanes separating different classes as well as possible. These areas were used to classify non-labelled individuals. A step-by-step algorithm was applied (Romeder, 1973).

Results

Odour recognition scores

Frequencies of individuals as a function of the number of correct responses are depicted in Figure 3 (top). All the subjects found at least four correct responses out of nine and 16 subjects obtained 100% of correct responses. The mean number of correct performances from 105 subjects was ~7.47 ± 1.08.

Frequencies of incorrect responses as a function of each different odour pair tested are given in Figure 3 (bottom). The highest score of incorrect responses was observed for the aldehyde C8–aldehyde C10 pair, and indeed it was close to a random answer (55 per 105).

Intensity ratings

Mean intensities are depicted in Figure 4 as a function of the pair (first factor) and of the place in each pair (second factor: first or second odour) when the recognition scores

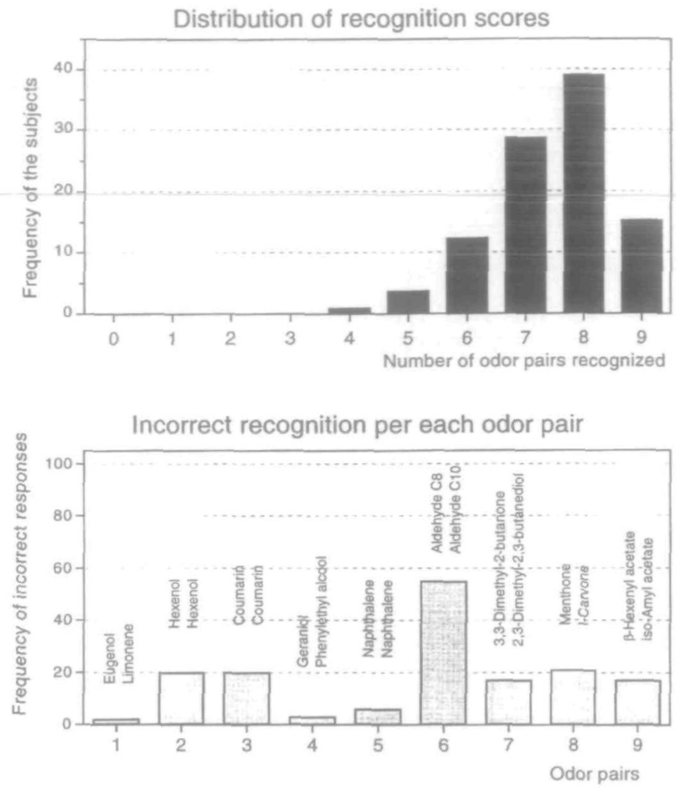

Figure 3 Frequencies of individuals as a function of the number of correct responses (top) and frequencies of incorrect responses (bottom) for each odour pair.

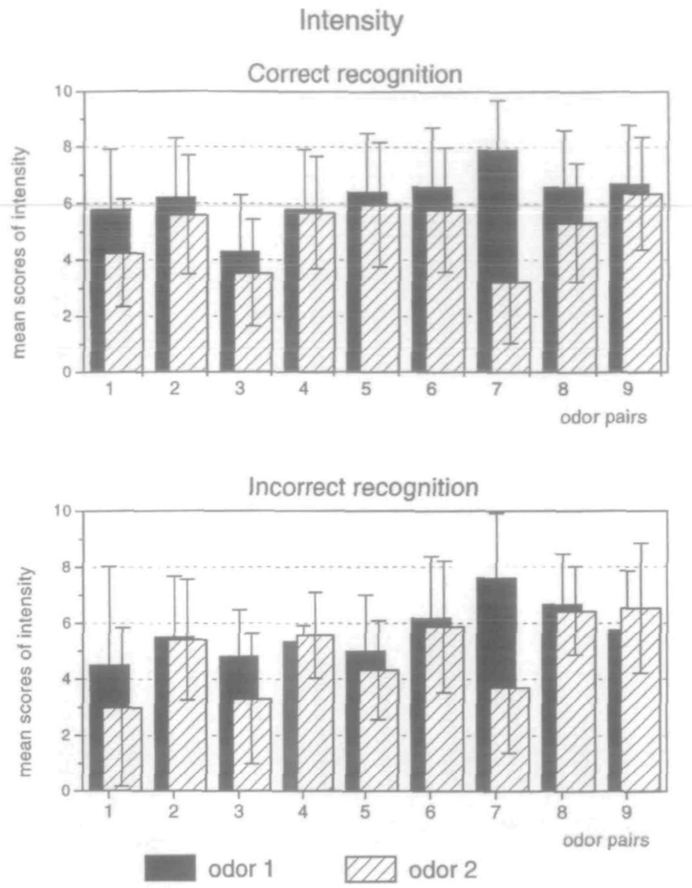

Figure 4 Mean scores of intensity as a function of the number of correct responses (top) and of the number of incorrect responses (bottom) for the first odour (black) and the second odour (hatched) of each pair. The vertical bars represent standard deviations.

were correct or incorrect. The analysis of data as a function of recognition scores does not allow a three-way analysis of variance (three-way ANOVA) to be performed to take this factor into account. The numbers of correct or incorrect responses were systematically different between the nine odour pairs while the measurements were repeated between the subjects for this condition. Therefore, two-way ANOVAs with repeated measurements on the second factor (Winer, 1962) were performed independently either when the recognition scores were correct or when they were incorrect.

When the recognition scores were correct, ANOVA showed significant differences of mean scores as a function of the pair condition [$F(8,775) = 16.03$, $P < 0.0005$], as a function of the place condition [$F(1,775) = 257.57$, $P < 0.0005$], and a significant interaction between these two factors [$F(8,775) = 42.93$, $P < 0.0005$]. Multiple orthogonal comparisons revealed significant decreases ($P < 0.05$ at least) of mean scores from the first to the second odour, except for pairs 4 (geraniol–phenylethyl alcohol) and 9 (β-hexenyl acetate–iso-amyl acetate). A very high decrease was noted for the seventh odour pair (3,3-dimethyl-2-butanone–2,3-dimethyl-2,3-butanediol) [$F(1,775) = 510.55$, $P < 0.0005$]. Thus, the differential intensity notes could play a role in discrimination when the two odours of a given pair were different. However, the same result was also obtained for the three identical odour pairs (pair 2: hexenol; pair 3: coumarin; pair 5: naphthalene). Therefore, the intensity note is not only a recognition factor, but also a false alarm factor.

When the recognition scores were incorrect, ANOVA did not show significant differences of mean scores as a function of the pair condition [$F(8,152) = 1.68$, NS], but a significant difference as a function of the place condition [$F(1,152) = 11.35$, $P < 0.0001$], and a significant interaction between these two factors [$F(8,152) = 3.87$, $P < 0.0005$]. Multiple orthogonal comparisons revealed significant decreases ($P < 0.05$ at least) of mean scores from the first to the second odour only for pairs 3 (coumarin–coumarin) and 7 (3,3-dimethyl-2-butanone–2,3-dimethyl-2,3-butanediol). However, it is noteworthy that the differences between the intensities of the two stimuli are either larger for pair 3 or

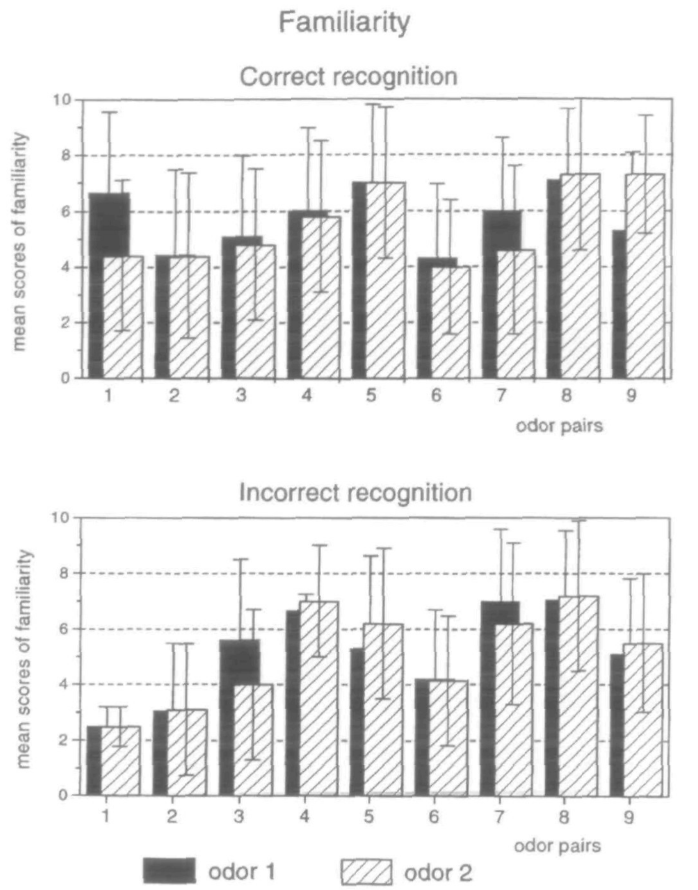

Figure 5 Mean scores of familiarity as a function of the number of correct responses (top) and of the number of incorrect responses (bottom) for the first odour (black) and the second odour (hatched) of each pair. The vertical bars represent standard deviations.

smaller for pair 7 than those respectively obtained when the recognition scores were correct. Thus, the error in evaluating the odour intensities seems higher when the recognition scores were incorrect than when they were correct.

Familiarity ratings

Mean familiarities are depicted in Figure 5 as a function of the pair (first factor) and of the place in each pair (second factor: first or second odour) when the recognition scores were correct or incorrect. As for intensity, two-way ANOVAs with repeated measurements on the second factor were performed either when the recognition scores were correct or when they were incorrect.

When the recognition scores were correct, ANOVA showed significant differences of mean scores as a function of the pair condition [$F(8,775) = 16.30$, $P < 0.0005$], as a function of the place condition [$F(1,775) = 6.76$, $P < 0.01$], and a significant interaction between these two factors [$F(8,775) = 15.41$, $P < 0.0005$]. Multiple orthogonal comparisons revealed significant decreases ($P < 0.0005$) of mean scores from the first to the second odour for pairs 1 (eugenol–limonene) [$F(1,775) = 71.05$, $P < 0.0005$] and 7 (3,3-dimethyl-2-butanone–2,3-dimethyl-2,3-butanediol) [$F(1,775) = 23.58$, $P < 0.0005$], and a significant increase for pair 9 (β-hexenyl acetate–iso-amyl acetate) [$F(1,775) = 49.38$, $P < 0.0005$]. Thus, these differential ratings of familiarity, as well as the absence of variations of familiarity scores for identical odour pairs (pair 1: hexenol; pair 2: coumarin; pair 5: naphthalene), could contribute to or partly explain the recognition performance of the subjects.

When the recognition scores were incorrect, ANOVA showed significant differences of mean scores as a function of the pair condition [$F(8,152) = 3.39$, $P < 0.005$], but no significant difference as a function of the place condition [$F(1,152) = 0.10$, NS], and no significant interaction between these two factors [$F(8,152) = 1.01$, NS]. Multiple orthogonal comparisons revealed significant decreases of mean scores from the first to the second odour for pair 3 (coumarin–coumarin) only [$F(1,152)=14.03$, $P < 0.0005$]. Three results deserve to be underlined relative to the correct recognition scores: a significant difference for a pair of identical odours (pair 3), no difference for two pairs of different odours (pair 7: 3,3-dimethyl-2-butanone–2,3-dimethyl-2,3-butanediol; pair 9: β-hexenyl acetate–iso-amyl acetate). These results were consistent with the incorrect recognition responses of the subjects, and could therefore account for their recognition performance.

Odour quality profiles

The odour descriptor profiles (137 semantic descriptors) of the nine odorants were determined as a function of the correct or incorrect recognition scores (Figure 6, solid line and grey columns respectively). Except for the aldehyde C8–aldehyde C10 pair, the number of the subjects incorrectly recognizing both odours in a pair was lower than the number of the subjects correctly recognizing them. Therefore, to compare both kinds of profiles as a function of the correctness of recognition visually, data were normalized to a percentage. Due to the very small number of incorrect scores for pairs 1 (eugenol–limonene), 4 (geraniol–phenyl ethyl alcohol) and 5 (naphthalene–naphthalene), we have represented the semantic profiles obtained only when the recognition scores were correct in Figure 6.

When the recognition scores were correct (Figure 6, solid line), most of the odorants could be characterized by a few discriminative main notes and a lot of secondary notes. For

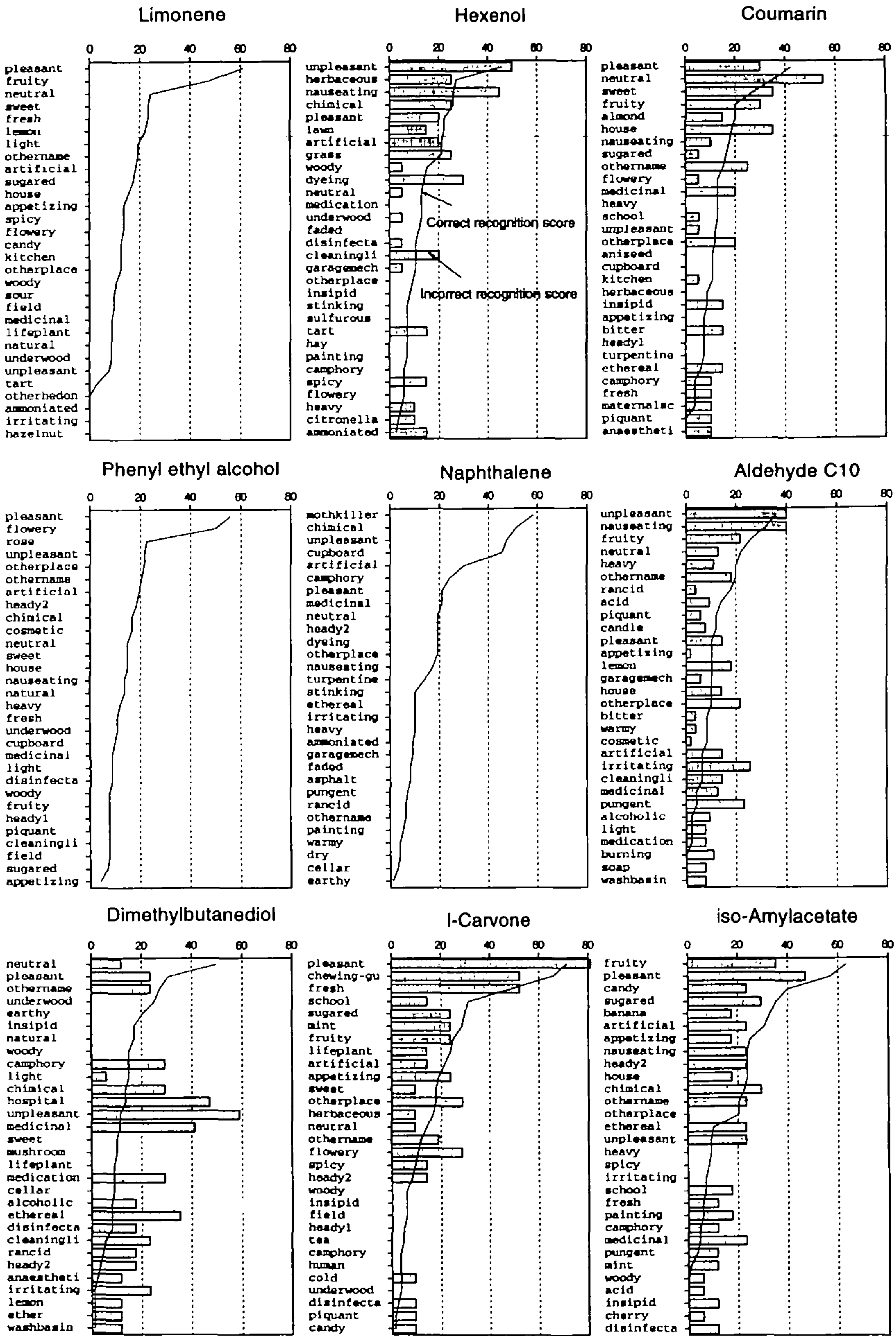

Figure 6 Odour quality profiles for the nine odorants as a function of the recognition scores. When the recognition scores were correct (solid line), the semantic profiles were ordered as a function of the decreasing number of descriptors (abscissa) selected by the subjects. When the recognition scores were incorrect (grey column), the semantic profiles were represented in the same order as previously. Only thirty descriptors are represented, indicating a high value and/or a high difference between both profiles.

instance, the odour of *l*-carvone (typically evoking the odour of a well-known aroma of chewing gum) was predominantly judged as pleasant, fresh and named as chewing gum. Qualities such as sugared, mint, school, fruity, plant could complete the odour impressions. Naphthalene and iso-amyl acetate also provided sufficiently discriminative descriptions. By contrast, aldehyde C10, coumarin and dimethylbutanedione did not give characteristic peaks except for hedonic descriptors such as pleasant, unpleasant or nauseating. Often, for these less well defined odorants, the mean odour profile indicates two or three descriptors selected simultaneously by the subjects and related to hedonic descriptors, such as pleasant and neutral for coumarin or unpleasant and nauseating for aldehyde C10. Thus, no clear profile or very slightly distinctive profiles were obtained for these odorants. However, when the correct (solid line) and incorrect (grey column) recognition judgements are compared, clear differences in the semantic profiles are observed. Thus, the frequency of each descriptor could be inferior, equal or superior to this of the reference profile (solid line). These variations were not given at random, because they were mainly limited to some descriptors from the 137 possible ones.

Memories

The frequencies of memories for each second stimulus are represented as a function of the recognition scores in Figure 7. As for the odour quality profiles, data were normalized in percentage and only the profiles obtained when the recognition scores were correct are represented for the pairs 1, 4 and 5. Naphthalene evoked more and limonene fewer memories than the other odours A chi-squared test for independence showed that the frequencies of memories for the six other odours are independent of the recognition scores [$\chi^2(1,5) = 8.62$, $P > 0.05$]. Thus, personal memories did not seem to influence the recognition performances differently, whatever the correctness of the subject's responses.

Connectionist approach

The performances obtained from the connectionist approach, when all data are taken into account, are depicted in Table 3 (complete data on left). The number of cells in this hidden layer varied from 3 to 33 according to the subjects' difficulty in discriminating between both odours of a pair and/or in selecting semantic descriptors. Performances for the learning and generalization phases were very high, except for the odour pairs 6 and 8: odours aldehyde C8–aldehyde C10 were found to be very hard to discriminate. Furthermore, the results indicated in Table 3 reveal that the connectionist approach systematically offered better predictive performances than the discriminant analysis for any odour.

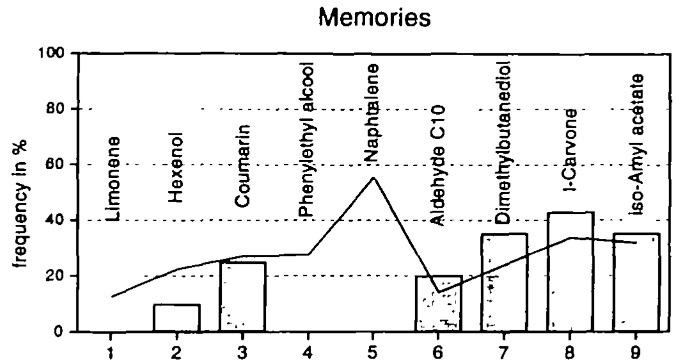

Figure 7 Frequency in percentage of memories for the nine odorants as a function of the correct (solid line) or incorrect (grey column) recognition scores Due to the very small number of errors for pairs 1, 4, and 5, only the profile obtained when the recognition scores were correct is represented.

To determine the relative importance of the descriptors in the olfactory performance of recognition memory, intensity, familiarity and hedonic criteria were suppressed. Thus, between 15 and 22 descriptors were submitted to new analyses (Table 3, reduced data). The number of cells in the hidden layer varied from 4 to 43. Learning and generalization performances were unchanged except for pairs 2 and 6. The absence of variation of learning performance can mean that intensity, familiarity and hedonic parameters did not play a preponderant role in olfactory recognition. Indeed, when nothing but these parameters were taken into account, the neural networks evidenced bad learning performances. They could not correctly learn to give correct responses because of insufficient information. For pair 2, the decrease of generalization performance (from 81 to 67%), when intensity, familiarity and hedonic scores were withdrawn from the data set, would mainly be due to hedonic descriptors. Thus, when only intensity and familiarity were withdrawn, the same rate of generalization of 81% as previously observed was obtained. Finally, for the odour pair 6, the surprising increase of the generalization rate from full data (62%) to reduced data (76%) could be due to contradictory data in the complete database: some examples have a different target output, but similar input data.

Table 3 Performances observed for each odour with the connectionist approach and the discriminant analysis

Odour pairs	Complete data					Reduced data		
	Connectionist approach			Discriminant analysis		Connectionist approach		
	Network a–b–c	Learning rate (%)	General rate (%)	Learning (%)	General (%)	Network a–b–c	Learning rate (%)	General rate (%)
1	28–3–1	100	95	–	–	17–4–1	100	95
2	35–20–1	99	81	61	43	22–5–1	98	67
3	28–17–1	99	81	70	57	17–15–1	95	81
4	29–7–1	100	100	90	95	16–12–1	99	95
5	30–8–1	99	95	72	66	18–4–1	99	95
6	32–33–1	100	62	80	56	20–43–1	94	76
7	29–7–1	100	86	65	52	17–15–1	98	81
8	28–16–1	99	76	69	54	19–10–1	98	76
9	27–6–1	99	81	72	45	15–18–1	98	81

Complete data: all data were used for analyses. Reduced data: intensity, familiarity and hedonic descriptors were suppressed from analyses.

a–b–c, Number of units on the input, hidden and output layers respectively.

–, calculations were not performed due to the too small number of misses obtained by subjects.

Discussion

Odour recognition scores

The distribution of odour recognition scores revealed the high mean performances of subjects since they correctly recognized odours as identical or different in 7.5 out of 9 pairs (83%). No subject produced more than three errors. The examination of recognition scores per each odour pair indicates few incorrect responses except for pair 6 (aldehyde C8–aldehyde C10). Jehl *et al.* (1994) tested olfactory recognition performances by using two sets of odour pairs: slightly dissimilar pairs and very dissimilar pairs. For a retention interval of 30 s, the scores obtained averaged 75% of correct responses for slightly dissimilar odour pairs (84% of hits for identical odour pairs and 64% of correct rejections), and 89% of correct responses for very dissimilar odour pairs (79% of hits for identical odour pairs and 99% of correct rejections). It was concluded that the percentage of correct rejections depends on the qualitative similarity of both odours in a pair. Thus, the high performances obtained in the present study can be explained by some pairs of odorants being easy to discriminate.

Intensity and familiarity scores

When the recognition scores were correct or incorrect, we have noted that the intensity note could be, respectively, not only a recognition factor, but also a factor leading to errors of judgement as false alarms. Thus, in the present study, intensity did not seem to be a decisive factor in explaining recognition, except perhaps for pair 7 (3,3-dimethyl-2-butanone-2,3-dimethyl-2,3-butanediol). In this case, a very large difference in intensity was observed while both odours were qualitatively different. In a previous paper, we also noted that intensity was not a prevailing factor of odour discrimination (Jehl *et al.*, 1994).

We have observed a tendency to perceive the intensity of the second odour with less acuity than the first one. Barker and Weaver (1983) have already observed that the remembered intensity is less than perceived intensity regardless of the retention interval. Apparently, it cannot be explained by an adaptation. Whereas the interval between both odours in a pair was 30 s in our study, Cain (1970) found that the magnitude estimates of suprathreshold intensity showed noticeable recovery after only three inhalations of fresh air. If we consider that the mean duration of an inspiration is 4 s, we can expect that the subjects recover from self- adaptation after 12 s. Under the same reasoning, the decrease of perceived intensity from the first to the second odour could not explained by cross-adaptation when the two odorants in a pair were different, whereas cross-adaptation has been shown to be almost always weaker than self-adaptation (Köster and De Wijk, 1991). In other respects, as previously indicated (Jehl *et al.*, 1994), it does not seem that this result can be explained by a quicker decline in the memory for subjective intensity than in the memory for qualitative impressions. No clear explanation can be given for interpreting the decrease of the intensity scores from the first to the second odour.

For familiarity, when the recognition scores were correct for different odour pairs, only three familiarity mean scores out of six varied from the first to the second odour, either by reduction (pairs 1 and 7), or by increase (pair 9). Thus, these differential ratings of familiarity, as well as the absence of variation of familiarity scores for identical odour pairs (pairs 1, 2 and 5), could contribute to or partly explain the recognition performance of the subjects. Moreover, when the recognition scores were incorrect, the results of familiarity scores could also be consistent with the incorrect recognition responses of the subjects. Thus, the subjects judging different two identical odours (pair 3) also rated their familiarity as different. Conversely, on average, they recognized as identical two different odours (pairs 7 and 9) and rated their familiarity as identical. Although giving erroneous responses, the subjects performed consistent recognition and familiarity judgements. We can conclude that familiarity could partly underlie recognition judgements.

Semantic profiles

Compared to the subjects who have correctly recognized a pair of odours as identical or different (i.e. the good performers), those who have incorrectly recognized them (i.e. the bad performers) have also described the second odour differently, but characteristically. As demonstrated here, a neural network can use these differential informations to predict the recognition performances of the subjects.

Two interpretations could be proposed to explain incorrect recognition. On the one hand, odours could be less well identified by the bad performers than by the good performers. For instance, the descriptors such as chewing gum, school, sugared for *l*-carvone, or fruity, candy, sugared, banana for iso-amyl acetate were less frequently used by the bad performers than the good ones. By contrast, they used less relevant descriptors such as chemical, medicinal, cleaning liquid, disinfectant, washbasin—cleaning liquid can represent a large class of substances with a variety of smells (lavender, citrus, solvent, turpentine, bleach, etc.). On the other hand, odours seemed to be perceived as less pleasant. Positive and negative hedonic descriptors (i.e. pleasant and appetizing versus unpleasant, nauseating, stinking and repugnant) were not used equally by the subjects to describe the nine odours, but their use varied as a function of the recognition judgement (correct or incorrect) (Table 4). First, more positive than negative hedonic descriptors were selected by the subjects when their recognition scores were correct. Second, negative hedonic descriptors were mainly chosen by the subjects when their recognition scores were incorrect. Third, the descriptor

Table 4 Frequency of the kind of hedonic descriptor such as pleasant (pleasant and appetizing), unpleasant (unpleasant, nauseating, stinking and repugnant) or neutral, as a function of the correctness of the recognition scores

Recognition score	Pleasantness		
	Pleasant	Unpleasant	Neutral
Correct	410	325	184
Incorrect	61	99	186

'neutral' was proportionally more frequently used when the recognition scores were incorrect than when they were correct. A chi-squared test for independence revealed that the frequency of the different types of hedonic descriptors are dependent of the recognition scores [$\chi^2(1,2) = 150.38$, $P < 0.005$]. Finally, regarding the other categories of descriptors, those presenting a negative hedonic connotation were more frequently used when the recognition scores were incorrect than when they were correct. For instance, for pair 7 (3,3-dimethyl-2-butanone–2,3-dimethyl-2,3-butane- diol), the frequencies of olfactory, gustatory, somesthesic, veridical label or place descriptors such as camphoraceous, chemical, hospital, medicinal, medication, ethereal, disinfectant, cleaning liquid, rancid, heady, anaesthetic, irritating, lemon, ether and washbasin were clearly higher, and those such as underwood, earthy, insipid, natural and woody were clearly lower.

In summary, the bad performers in recognition task used unspecific terms such as 'cleaning liquid' to qualify the odorous. They also seemed to perceive the odours as being rather unpleasant and indeed used many descriptors presenting a negative hedonic connotation.

Mean profile and olfactory memory recognition

As reported above, this experimental study suggests that the accuracy of odour descriptions and the correctness of recognition scores are related. Furthermore, this study also shows by simulation that it is possible to predict the olfactory short-term recognition performance of an individual from his/her odour characterization profile. Depending on the odour pair, the neural networks could predict, with success ranging from 62 to 100%, which subjects could correctly determine that two odours in a pair are identical or different. Thus, the network was able to extract relevant information from evocations such as familiarity or more specific semantic descriptors. Individuals who provided semantic profiles similar to the reference mean profile of the database, i.e. who 'correctly' described odours, also correctly determined odour similarities or differences in the short-term memory recognition task. By contrast, individuals whose odour descriptions differed from the reference mean profiles, recognized the odour pairs badly. This suggests the existence of mean profiles to describe odours, on the one hand, and, on the other hand, of a relationship between semantic and perceptual characteristics.

The use of odour profiling is well established method for characterizing odorants. From a given population and from rating-scale information obtained with a set of descriptors, the experimenter can deduce a mean odour profile. Unless experts are requested for studies, a number of subjects are needed to mitigate the problem of bad descriptions and to present a mean profile. The present study suggests that this mean odour profile can also be obtained by distinguishing good and bad performers in an olfactory recognition memory task. Thus, on the one hand, a subject's memory olfactory performance can be predicted by the profiling of the memory of those odorants—primarily because both processes are based on the same initial stimuli and resultant memory trace. But also, on the other hand, the distinction of subjects as a function of their recognition performances allows two types of mean profiles to be distinguished: a profile with specific descriptions and a profile with unspecific terms.

Previous observations have shown that retention improves if an item is encoded in terms of semantic characteristics rather than only perceptual characteristics. This effect was observed in the framework of olfactory recognition memory studies when odours were paired with a short delay and had to be discriminated (Rabin and Cain, 1984; Walk and Johns, 1984; Lyman and McDaniel, 1986; Jehl *et al.*, 1996). For instance, Lyman and McDaniel (1986) observed that subjects instructed to associate a label-plus-definition with the target odours of common food substances exhibited a higher memory performance than control subjects who were instructed to smell each odour only. Jehl *et al.* (1996) tested short-term or long-term odour recognition memory performances of subjects who had previously learned either to associate a 'veridical' label, or a generated object name or a chemical name with each odorant. The results showed that discrimination performances were increased for 'veridical' names in short-term memory and for generated and veridical conditions in long-term memory. The more accurate the label, the more likely the odour was correctly recognized. Performances with chemical names did not differ significantly from no-labelling conditions when odours were previously presented without verbal associations.

Various processing levels?

The present results suggest experimentally, but also via a simulation, that a good performer in a recognition task also gives a more accurate description of the odours than a bad performer. These results also suggest that all the different types of descriptors did not play a similar role in the recognition performance. Thus, according to the present results, intensity did not seem to contribute significantly to odour recognition. Also, by simulation, we have shown that, together, the intensity, familiarity and hedonic judgements did not give the neural network pertinent information allowing it to learn how to predict which subjects correctly discriminated between both odours of a pair. These data would not allow the subject to recognize the odours unambiguously, because they are not, or only a few of them are, discriminating odours. By contrast, the information contained in the olfactory, gustatory and visual associations evidenced by the semantic descriptors would be more specific and, consequently, more discriminatory for the neural network. In this respect, the simulation seems to mimic the actual olfactory processing. Intensity, familiarity and pleasantness judgements can be regarded as lower-level processing issues, while the olfactory, gustatory and visual associations producing semantic descriptions would represent higher-level processings. Schab (1991) suggested that the process of olfactory identification comprises different levels of analyses with performances ranging from non-verbal feelings of familiarity to specific object names. In a more general frame, Craik and Lockhart (1972) proposed that incoming sensory stimuli can be analysed at different levels of processing. For instance, visual or auditory stimuli would range from shallow processing, such as form, pitch or colour, to deeper semantic processings. Furthermore, Craik and Tulving (1975) demonstrated that correct recognition of verbal stimuli depends on the type of encoding that the subject was induced to perform. This dependency of memory performance on different types of encoding operations was termed the 'levels of processing effect'. Further studies are needed to determine if similar levels of processing do exist in olfaction. If we suppose that the functional duality between low- and high-level processings is supported by separate nervous mechanisms, cerebral imaging techniques could help us to distinguish successfully between semantic and perceptual representations of the odours.

ACKNOWLEDGEMENTS

This study was supported by the Cognitive Sciences Programme of Centre National de la Recherche Scientifique, France.

REFERENCES

Barker, L.M. and Weaver, C.A. (1983) Rapid, permanent, loss of memory for absolute intensity of taste and smell. *Bull. Psychon. Soc.*, **21**, 281–284.

Brown, R.W. and McNeill, D. (1966) The 'tip of the tongue' phenomenon. *J. Verbal Learn. Verbal Behav.*, **5**, 325–337.

Cain, W.S. (1970) Odour intensity after self-adaptation and cross-adaptation. *Percept. Psychophys.*, **7**, 272–275.

Craik, F.I.M. and Lockhart, R.S. (1972) Levels of processing: a framework for memory research. *J. Verbal Learn. Verbal Behav.*, **11**, 671–684.

Craik, F.I.M. and Tulving, E. (1975) Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.*, **104**, 268–294.

Dayhoff, J. (1990) *Neural Network Architectures: An Introduction*. Van Nostrand Reinhold, New York.

Doty, R.L. (1991) Olfactory system. 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. 175–203.

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

Fisher, R.A. (1936) The use of multiple measurements in taxonomic problem. *Ann. Eugen.*, **7**, 179–188.

Jehl, C., Royet, J.-P. and Holley, A. (1994) Very short term recognition memory for odours. *Percept. Psychophys.*, **56**, 658–668.

Jehl, C., Royet, J.-P. and Holley, A. (1995) Odour discrimination and recognition memory as a function of familiarization. *Percept. Psychophys.*, **57**, 1002–1011.

Jehl, C., Royet, J.-P. and Holley, A. (1996) Role of verbal encoding in olfactory memory. *Percept. Psychophys.*, in press.

Kōster, 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*. Springer-Verlag, Berlin, pp. 199–215.

Lawless, H.T. and Engen, T. (1977) Associations to odours: interference, mnemonics, and verbal labeling. *J. Exp. Psychol., Hum. Learn. Mem.*, **3**, 52–59.

LeCun, Y. (1985) Une procédure d'apprentissage pour réseau à seuil asymétrique. *Proc. Cognitiva 85*. Paris, pp. 599–604.

Lyman, B.J. and McDaniel, M.A. (1986) Effects of encoding strategy on long-term memory for odours. *Q. J. Exp. Psychol.*, **38A**, 753–765.

Moskowitz, H.R. (1974) Combination rules for judgments of odour quality difference. *Agric. Food Chem.*, **22**, 740–743.

Moskowitz, H.R. and Barbe, C.D. (1976) Psychometric analysis of food aromas by profiling and multidimensional scaling. *J. Food Sci.*, **41**, 567–571.

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

Paugam-Moisy, H. (1991) Selecting and parallelizing neural networks for improving performances. In Kohonen, T., Makasira, K., Simula, O. and Kangas, J. (eds), *Artificial Neural Networks*. North-Holland–Elsevier, Amsterdam, pp. 659–664.

Rabin, M.D. and Cain, W.S. (1984) Odour recognition: familiarity, identifiability, and encoding consistency. *J. Exp. Psychol.: Learn. Mem. Cogn.*, **10**, 316–325.

Romeder, J.M. (1973) *Méthodes et Programmes d'Analyse Discriminante*. Dunod, Paris.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) Learning internal representations by error propagation. In Rumelhart, D.E., McClelland, J.L. and the PDP Research Group (eds), *Parallel Distributed Processing*. MIT Press, Cambridge, MA, pp. 318–362.

Schab, F.R. (1991) Odour memory: taking stock. *Psychol. Bull.*, **109**, 242–251.

Tukey, J. (1958) Bias and confidence in not quite large samples. *Ann. Math. Statist.*, **29**, 64.

Walk, H.A. and Johns, E.E. (1984) Interference and facilitation effects in short-term memory for odours. *Percept. Psychophys.*, **36**, 508–514.

Winer, B.J. (1962) *Statistical Principles in Experimental Design*. McGraw-Hill, New York.

Received on November 29, 1995; accepted on July 15, 1996