Effects of Panel Experience on Olfactory Memory Performance: Influence of Stimuli Familiarity and Labeling Ability of Subjects

Isabelle Lesschaeve and Sylvie Issanchou

INRA, Laboratoire de Recherches sur les Arômes, 17 rue Sully, B.V. 1540 21034 Dijon cedex, France

Correspondence to be sent to: I. Lesschaeve, INRA, Laboratoire de Recherches sur les Arômes, 17 rue Sully, B.V. 1540 21034 Dijon cedex, France

Abstract

This work attempted to define the impact of panel experience on olfactory memory performance by comparing scores in an odor recognition task obtained from a highly trained descriptive panel (17 subjects) and a naive one (33 subjects with no experience in sensory analysis). During the inspection phase, 16 odorants were presented monadically to subjects for familiarity rating and a written description. The recognition session was planned 7 days later with 32 odorants (including the 16 of the target set). Subjects also described the odor of the stimuli. The memory performance of each panel was estimated by the mean value of individual *d'* (index of detectability). Training of the descriptive panel did not enhance the olfactory memory ability, which strengthened the well-established idea that odor memory is closely linked to personal experience of the subject with the stimuli. However, odor recognition performances were improved by a precise and consistent answer at the encoding task: for both groups, subjects performed better when stimuli were precisely labeled and consistently used. Trained subjects were more accurate when describing than naive ones, whatever the degree of familiarity. Panel experience allowed trained subjects to verbalize their perception especially accurately when stimuli were familiar. **Chem. Senses 21: 699–709, 1996.**

Introduction

Sensory profiling is a sensory method which requires the use of human subjects as the measurement tool. These subjects are selected on specific abilities, e.g. no olfactory or gustatory deficiencies, discrimination and verbalization, as described by many authors in the literature (Amerine *et al.*, 1965; ASTM, 1981; Stone and Sidel, 1985; Meilgaard *et al.*, 1991; ISO, 1993). Descriptive panelists are then encouraged to develop a product-specific vocabulary to describe all the different sensory attributes they perceive in the range of studied products. When the descriptive item list is defined, subjects have to be trained to identify the sensory perception in the product and to quantify it, in order to obtain consistent sensory data, in terms of product differences (Stone *et al.*, 1974). By applying this process to flavor characteristics, panelists have to learn to associate a specific label with a specific olfactory perception. When the number of descriptors is >15, this task seems closely linked to odor recognition ability, and to a larger extent to odor memory performance. Indeed, we can hypothesize that an efficient descriptive panelist will also produce good performances on odor memory and recognition tests. As far as we know, the literature in that field has never investigated the effect of such training on these particular performances. Desor and Beauchamp (1974) observed an increase of odor recognition performances when subjects were especially trained on the tested stimuli. In that case, odorants became more and more familiar to the subjects, who consequently improved their scores. The influence of familiarity on odor memory and odor recognition performances was later demonstrated by Rabin and Cain (1984). Moreover, familiarity with the target stimuli gained by training was found to also enhance performances in an odor quality discrimination task (Rabin, 1988) in which a memory-trace component is also involved. However, experience in descriptive sensory analysis has never been studied as a potential influencing factor.

The aim of this work was to define the global impact of a training program for sensory profiling on odor recognition performance by comparing data obtained from an experienced descriptive panel and a naive one. Moreover, the influence of familiarity, labeling precision and label use consistency on this ability are also examined.

Materials and methods

Subjects

Experienced subjects (17 women, age range: 17–60 years old) had been selected as previously described (Issanchou *et al.*, 1995). They had been trained to perform a conventional sensory profile on Camembert cheeses during 33 sessions (~50 h). During the first four training sessions, subjects had learnt to analyze texture and flavor perceptions in different media. Then the descriptive vocabulary was generated by tasting 15 different commercial Camembert cheeses (five sessions). Subjects then learnt to evaluate cheeses according to the profiling technique by tasting 27 commercial samples (nine sessions): these sessions permitted vocabulary alignment amongst the panelists by defining each descriptive term with a flavor standard or a verbal definition. The appropriateness of the attributes for describing the experimental Camembert cheeses had been checked thereafter. At the end of this intensive training, subjects had been employed as panelists during 12 additional sessions (12 h) to evaluate the experimental samples. They had never been exposed to the stimuli used in the present experiment during all these sessions.

Naive subjects (2 men, 31 women, age range: 17–55 years old) were recruited from our consumer database. They had no previous panel experience and were matched for age with the experienced panel.

Stimuli

Thirty-two odorant solutions were prepared from either food product, essential oils or volatile compounds. Moreover, the experimenter chose them in order to have, *a priori*, 16 familiar odors and 16 less familiar. They are listed in Table 1. Twenty-five milliliters of each odorant solution were presented in 60 ml brown coded glass flasks.

Experimental conditions

All sessions took place in a sensory room equipped according to the recommendations of AFNOR (1987). Subjects evaluated the samples in separated booths; the room temperature was controlled (20 ± 1°C). The experienced panel was split into two groups; the naive one was split in five groups. This partition was due to the subjects' availability and was not a factor in our experimental design.

Procedure

Two sessions were organized, 7 days apart. The first session was the inspection of the odorant solutions. Sixteen samples were presented in the same randomized order to all the subjects; indeed, we were more interested in individual performance differences than in the comparison of compounds. For each solution, subjects were asked to rate their familiarity with the perceived odor on a 100 mm linear scale (left anchor: unfamiliar; right anchor: very familiar). Thereafter, they were asked to describe the perceived odor as precisely as they could. Thirty-two odorants were examined during the second session, the recognition phase. Half of the stimuli were the same as in the inspection phase (targets) and half were new (distractors). Subjects were simply told that amongst the samples, some were examined last week and some were new. The 32 solutions were presented in the same randomized order to all subjects. For each one, they had to say if the solution was presented during the first session (yes/no). Then they were asked to describe the perceived odor as precisely as they could.

For both sessions, subjects were asked to take a 30 s break between each sample to avoid adaptation (Engen and Ross, 1973) and not to return to their previous answers or to previous samples in order to preserve spontaneity of their answers.

The total time taken to perform each test was recorded for each subject.

Table 1 Stimuli

Compound	Code	Origin	SS (p.p.m.)	FS (p.p.m.)
Tarragon	tarr	essential oil[a]	5000	50
Thyme	thym	essential oil[a]	5000	50
Anethol	anet	[i]	1000	20
Geraniol	gera	[i]	1000	5
Menthol	ment	[i]	1000	30
γ-Nonalactone	nona	[i]	1000	20
Camphor	camp	[i]	2000	20
Strawberry	stra	[b]		10
Coriander	cori	essential oil[a]	5000	50
Viandox®	vian	commercial		10
Limonene	limo	[i]	1000	20
Phenyl acetaldehyde	phac	[i]	1000	1
Octanoic acid	octa	[i]	1000	50
(Z)-jasmone	jasm	[i]	1000	30
4-Ethyl phenol	4eph	[i]	1500	15
Hop	hop	essential oil[c]	1000	1
Black pepper		essential oil[d]	1000	20
Apple		natural flavor[e]	1000	8
1-Decanol		[i]	5000	25
Phenol		[i]	10000	120
α-Pinene		[i]	1000	5
δ-Decalactone		[i]	2000	20
Raspberry		[f]	50000	250
4-Methyl acetophenone		[i]	1000	10
Savory		essential oil[a]	5000	50
Saffron		essential oil[g]		10[h]
6-Methyl-5-hepten-2-one		[i]	1000	2
Guaiacol		[i]	1000	10
Ethyl 2-methylbutanoate		[i]	1000	1
Linalyl acetate		[i]	1000	5
Hexyl acetate		[i]	1000	5
Chicory		Leroux		40g/l

Compounds in bold letters belong to the target set; others are the distractants. SS: stock solution FS: final solution.

[a] CRMM (France); [b] Robertet (France); [c] Kalsec (USA); [d]: made in the laboratory; [e] given by Dr Petro-Turza (CFRI, Hungary); [f] Firmenich (Switzerland); [g] liquid spices Amora (France); [h] in paraffin oil (Fabioleol); [i] pure chemical compounds from the laboratory library.

Data analysis

Recognition performance was determined according to Signal Detection Theory (Banks, 1970), i.e. by computing hit and false alarm rates and d' scores. The index d' represents the index of detectability and it is obtained by the following formula (see Engen, 1971) :

$$d'_i = Z_{hits} - Z_{false\_alarms}, \text{ for a subject } i,$$

where Z_{hits} is the *z*-transformed percentage of effective recognitions and Z_{false_alarms} is the *z*-transformed percentage of false alarms.

Groups' d's were assessed by averaging individual d'_is as previously recommended by Rabin and Cain (1984).

Familiarity scores were obtained by measuring the distance of the subject's mark from the left anchor (score = 0). [The values are ranged from 0 (unfamiliar) to 100 (very familiar).]

As some stimuli did not have a well defined associated label (especially for some volatile compounds), the generated label was not analyzed versus a veridical name of the odor. However, the precision of the term given at the inspection phase and the consistency of its use during the recognition phase were recorded.

The first label precision was evaluated independently by two persons. When a subject gave a specific noun (e.g. ancient rose; vanilla; cowshed) the descriptor was considered as precise (and it was coded PN). When the odor corresponded to an expected categorization of the stimulus (e.g. red berries or fruity for strawberry; citrus for limonene) the descriptor was assessed as belonging to an odor family (coded OF). Finally, when the description was imprecise (e.g. women's perfume, aggressive odor, spring-like smell), it was assessed as a vague descriptor (coded VD).

Label use consistency was assessed independently by the same two people. Responses were coded following the most described scheme (Rabin and Cain, 1984; Lyman and McDaniel, 1986): CO for consistent label (i.e. exactly the same as in the first session); NM for near miss labeling (e.g. label 1: strawberry, label 2: raspberry; or label 1: rose, label 2: floral); and FM for far miss labeling (e.g. label 1: vanilla, label 2: clove).

When the two evaluaters disagreed on the evaluation of precision or consistency, they discussed the matter to reach a consensus.

All analyses and graphical representations were conducted by using SAS (SAS, 1988a,b). Two-tailed unpaired *t*-tests

Table 2 Odor recognition results

Group	n	% hits	% false_alarms	Averaged d'_i (SD; min; max)
Naive	33	62.1	24.6	1.23 (0.93; 0.16; 4.39)
Experienced	17	64.3	26.8	1.18 (0.88;-0.17; 4.03)

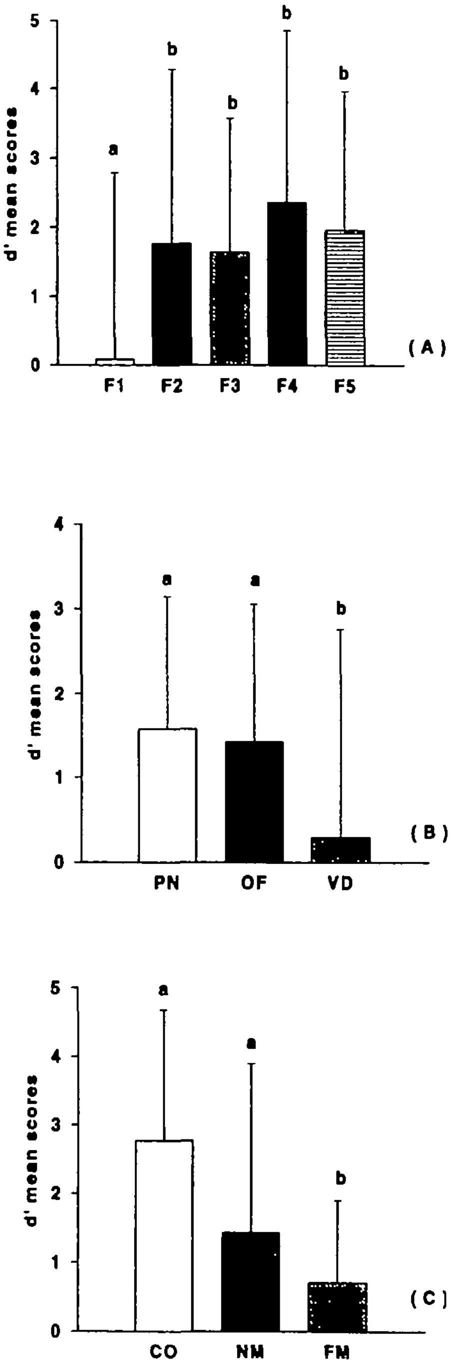

Figure 1 Influence of familiarity rating categories **(A)**, label precision categories **(B)** and label use consistency categories **(C)** on d' mean scores of both panels. The vertical bar represents the mean value, the line on each bar is the SD ($P < 0.05$). Bars with the same letters are not significantly different (least squares means t-test, $P < 0.05$).

were used to compare performances of each group. Analysis of variance (PROC GLM of SAS) was performed between and within groups to test the effects of familiarity, first label precision or label use consistency on performance. In case of a significant difference, a comparison of means was performed using a t-test computed on the least squares means (as the number of observations could be different in each class).

Results

d' and panel experience

The average d's of each group are reported in Table 2. The two values are very similar, as are their SD and range. The t-test did not reveal any significant difference between them [$t(48) = -0.18$; $P = 0.85$]. Moreover, there were also no obvious differences in the proportions of hits and false alarms between the two panels.

d' and familiarity

Odor memory performances are improved when stimuli are familiar to subjects (Rabin and Cain, 1984). To determine whether this assumption was verified by our data, familiarity scores were categorized in five classes, defined according to the mean scores distribution of the 16 target stimuli:

F1 unknown odor, score <10
F2 $10 \leq$ score < 30
F3 $30 \leq$ score < 60
F4 $60 \leq$ score < 75
F5 well known odor, score ≥ 75

d'_is were calculated within each familiarity category by assessing the percentage of hits in the category and the percentage of false alarms for all distractor odorants (as previously described by Rabin and Cain, 1984). An analysis of variance was performed according to the mixed model:

$$d'_i = \text{familiarity category} + \text{group} + \text{subject(group)} + \text{group*familiarity category} + \varepsilon$$

Results showed no average difference between groups [$F(1,48) = 0$; $P = 0.99$]. However, familiarity categories did affect d'_i values [$F(4,168) = 7.37$; $P = 0.0001$]. The least squares means comparisons showed that unknown odors (F1) led to lower d'_is. As no significant interaction was observed between group and familiarity category [$F(4,168) = 0.98$; $P = 0.42$], we concluded that the performances of the two groups moved in the same direction according to the odorant familiarity. Therefore, average results are reported in Figure 1A.

d′ and first label precision

The assessment of the precision could have three values: PN, OF or VD. Indices of d'_i were calculated within each category by assessing the percentage of hits in the category and the percentage of false alarms for all distractor odorants. An analysis of variance was performed according to the mixed model:

$$d'_i = \text{label precision category} + \text{group} + \text{subject(group)} + \text{group}^{*}\ \text{label precision category} + \varepsilon$$

The performances of the two groups were found to be similar [group effect, $F(1,48) = 0.05$; $P = 0.82$]. Although label precision did affect d'_i values significantly [$F(2,83) = 4.69$; $P = 0.01$], this effect did not vary significantly according to the groups, as shown by the non-significant interaction [$F(2,83) = 0.03$; $P = 0.97$]. Therefore the average results are reported in Figure 1B. The least squares means comparisons revealed that VDs induced significant lower d'_i values.

d′ and label use consistency

Label use consistency scores were categorized as described above in three classes: CO, NM and FM. Indices of d'_i were calculated within each category by assessing the percentage of hits in the category and the percentage of false alarms for all distractor odorants. An analysis of variance was performed according to the mixed model:

$$d'_i = \text{consistency category} + \text{group} + \text{subject(group)} + \text{group}^{*}\ \text{consistency category} + \varepsilon$$

Once again, the analysis revealed no group effect [$F(1,48) = 0.58$; $P = 0.44$] nor any interaction [$F(2,94) = 0.38$; $P = 0.68$]. However, consistency categories did significantly affect d'_i values [$F(2,94) = 1.52$; $P = 0.0001$]. Therefore, means d'_is for each consistency category are reported in Figure 1C. The least squares means comparisons showed that FM labels led to lower d'_i values.

Sensory experience did not appear as a determinant factor influencing recognition memory performances. Contrary to our hypothesis, experienced subjects who should have been better able to name odors did not perform better than naive subjects.

Target set evaluation

The set of familiarity degrees was chosen to be equivalent to both panels. Table 3 shows familarity ratings for both panels and each compound. The 16 compounds were ranked in approximately the same order by both panels. The familiarity range of tested odors was similar for both groups and thus the chance to have good performances was also equivalent. The most familiar stimulus was strawberry and the least familiar was hop. However, camphor, menthol and 4-ethyl phenol (to a lesser extent) appeared more familiar to the experienced subjects than to the naive ones. This was confirmed by unpaired t-tests (Table 3).

Experienced panelists were expected to have a better ability for labeling odorants consistently. This could be checked through the scores obtained, for all 16 odorants, for the precision of the first label and for the consistency of its use. Mean proportions of assessments for each precision category and each consistency of use were determined for each panel in order to compare scores obtained in each case. Results are reported on Tables 4 and 5. The effect of descriptive experience is noticeable, i.e. experienced panelists used less vague descriptors than did naive subjects. Moreover, experienced subjects tended to use more PNs ($P = 0.06$). They also used generated labels more consistently than did naive ones. The latter ones used more FM labels on the second presentation of the stimuli. By cross-matching the two criteria (see Table 6), we observed that a PN did not lead systematically to a consistent use. It actually occurred for 51 and 47% of stimuli judgements made by experienced and naive subjects respectively; a higher percentage was expected for the trained panel. This lack of consistency could be related to the low degree of familiarity with the stimuli. In fact, the encoding by a simple labeling task of a new stimulus (especially chemical compounds) seemed inefficient for a good retrieval (Cain, 1979). However, when the first label was imprecise (VD), the experienced panel was slightly more consistent than naives because 28% of these labels were named as NM; this occurred for only 6% of the

Table 3 Familiarity ratings

Compound[a]	Experienced		Naives		Student's *t*	Associated probability
	Familiarity[b]	SD	Familiarity	SD		
Hop	25.94	29.23	28.00	22.97	–0.27	0.78
Octanoic acid	28.82	28.43	29.42	23.03	–0.08	0.93
(Z)-Jasmone	36.29	35.47	39.97	29.96	–0.38	0.70
Tarragon	38.24	35.76	47.15	32.20	–0.89	0.37
Limonene	44.18	32.78	47.52	26.91	–0.38	0.70
Phenyl acetaldehyde	48.59	27.07	45.30	26.91	0.40	0.68
Geraniol	53.53	33.62	42.58	30.40	1.16	0.25
Coriander	54.00	29.79	46.61	30.00	0.82	0.41
4-Ethyl phenol	**54.47**	31.97	**37.03**	26.08	2.07	0.04
Thyme	60.29	31.80	58.24	25.05	0.25	0.80
Anethol	62.65	24.06	59.76	27.70	0.36	0.71
γ-Nonalactone	68.13	26.94	72.30	23.14	–0.56	0.57
Viandox®	71.88	26.86	65.45	25.26	0.83	0.40
Menthol	**74.41**	23.84	**55.48**	29.20	2.30	0.02
Camphor	**79.47**	22.78	**63.67**	22.37	2.36	0.02
Strawberry	86.41	14.38	81.27	18.86	0.98	0.33

Values in bold letters on the same row are significantly different between experienced and naive panelists (Student's *t*, $P < 0.05$).
[a] Compounds are ranked in ascending order according to the average of familiarity scores of the experienced panel.
[b] Familiarity scores expressed in mm (min = 0; max = 100).

Table 4 Between group comparison on first label precision

Precision assessment	Group	*n*	Mean proportion (%)	SD	Student's *t*	Associated probability
Precise noun	Ex	17	52.21	14.65	1.90	0.063
	Na	33	42.80	17.40		
Odor family	Ex	17	40.07	10.38	0.12	0.90
	Na	33	39.58	14.87		
Vague descriptor	Ex	17	7.72	7.82	–3.15	0.003
	Na	33	17.61	14.36		

Ex: experienced panel; Na: naive panel.

VDs used by the naives. The general descriptive training seems to improve the ability of subjects to categorize VDs in such a way that they could retrieve it approximately at the second presentation. A contingency table was obtained by cross-matching label precision and familiarity categories. The influence of familiarity on the label precision (Table 7) was particularly noticeable for experienced panelists. Precision was higher for that group when stimuli were familiar (69 versus 53% for naives). These conclusions could be checked for each of the 16 stimuli. Figure 2 shows the percentage of each label precision category obtained by stimulus and for each group. Stimuli were ranked in approximately the same order according to the average of the experienced panel's familiarity ratings (less familiar odorant on the left, most familiar one on the right); this choice was arbitrary and allowed the visualization of the influence of this factor on label precision. Although influence of familiarity on label precision was generally

Table 5 Between group comparison on label use consistency

Label use assessment	Group	*n*	Mean proportion (%)	SD	Student's *t*	Associated probability
Consistent	Ex	17	38.60	10.41	2.23	0.0355
	Na	33	29.73	14.49		
Near miss	Ex	17	23.53	8.71	1.68	0.0985
	Na	33	17.80	12.51		
Far miss	Ex	17	37.87	10.24	−3.61	0.0007
	Na	33	52.46	18.28		

Ex: experienced panel; Na: naive panel.

observed, there were some exceptions; for example, octanoic acid was judged unfamiliar by each panel, but 70% of experienced panelists gave a PN for it (versus 52% of the naives). Figure 3 shows the percentage of each label use consistency category obtained by stimulus and for each group. Stimuli are ranked on the same principle. The greater ability of experienced panelists to use consistent labels was verified on most of the tested stimuli except for tarragon, (Z)-jasmone and, to a lesser extent, 4-ethyl phenol. In fact, tarragon was more familiar to naives, which led to a consistent labeling.

The sensory profiling training was also noticeable in the time taken to perform the test. Experienced subjects needed on average 16 min for the inspection phase and 24 min for the recognition task. Naives needed 26 and 35 min respectively. A *t*-test allowed us to conclude that these durations were significantly different between the groups for both the inspection ($t = 5.88$, $P = 0.0001$) and recognition sessions ($t = 4.33$, $P = 0.0001$).

Discussion

Odor recognition performance (expressed by d') is not related to previous sensory profiling experience. Although experienced panelists had much practice in odorant description and sensory profiling, they did not perform better than the naives, which means that sensory profiling training had no positive influence on d' score. However, the d' values were very close to those obtained by Lyman and McDaniel (1986) for an equivalent encoding task (label plus definition, d' scores = 1.21), but slightly lower than those obtained by Rabin and Cain (1984) in the same conditions (d' scores = 1.50).

Table 6 Influence of label precision on label use consistency: mean proportion of individual answers within each category for each panel

Precision	Consistency (%)					
	Experienced			Naive		
	CO	NM	FM	CO	NM	FM
PN	50.70	21.83	27.46	47.35	17.26	35.40
OF	30.28	24.77	44.95	22.97	23.44	53.59
VD	0.0	28.57	71.43	2.15	6.45	91.40

Mean proportions were computed as follows: number of responses within a cross-category (precision*consistency)/total number of responses within the precision category. CO: consistent; NM: near miss; FM: far miss; PN: precise name; OF: odor family; VD: vague descriptor.

Table 7 Influence of familiarity on label precision: mean proportion of individual answers within each category for each panel

Familiarity	Precision (%)					
	Experienced			Naives		
	PN	OF	VD	PN	OF	VD
F1	24.32	54.05	21.62	31.94	31.94	36.11
F2	38.64	50.00	11.36	29.63	46.91	23.46
F3	47.37	42.11	10.53	44.14	40.00	15.86
F4	58.06	38.71	3.23	44.19	41.86	13.95
F5	68.93	30.10	0.97	53.47	37.50	9.03

Proportions were computed as follows: number of responses within a cross-category (familiarity*precision)/total number of responses within the familiarity category. F1: unknown odor, score <10; F2: $10 \leq$ score < 30 ; F3: $30 \leq$ score < 60 ; F4: $60 \leq$ score < 75; F5: well known odor, score ≥ 75. PN: precise name; OF: odor family ; VD: vague descriptor.

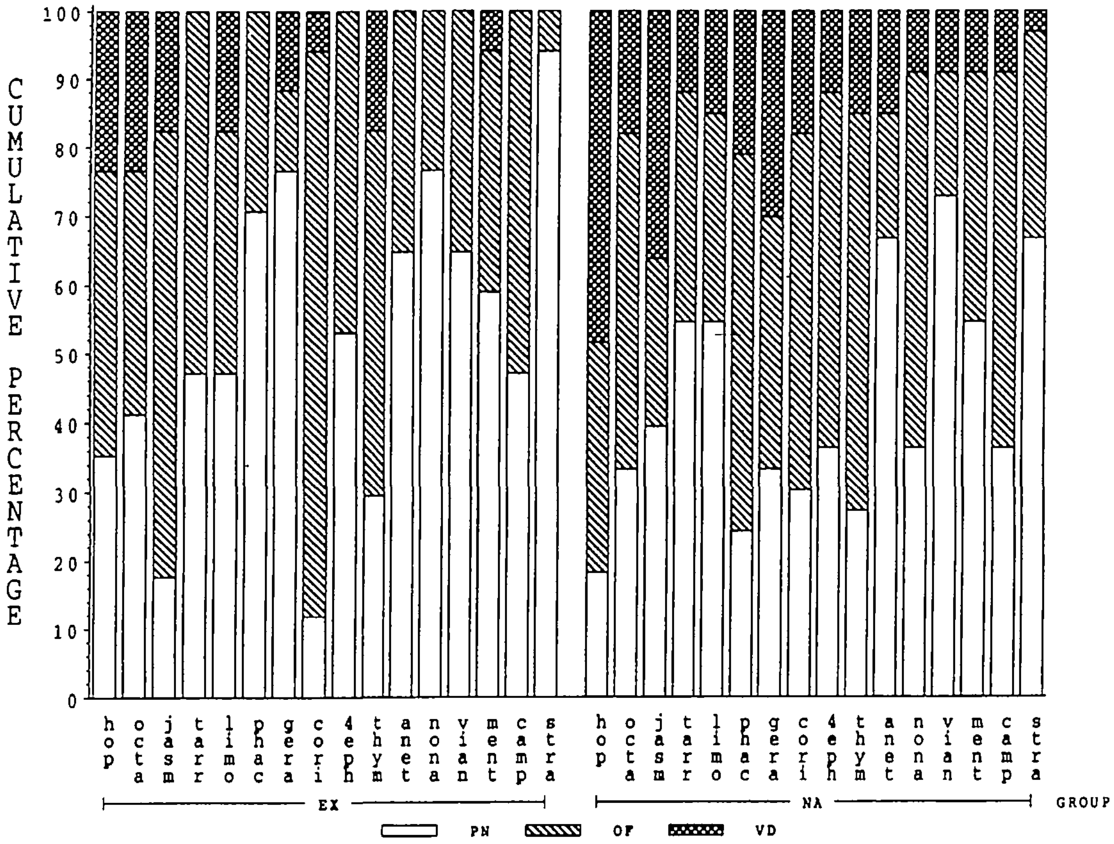

Figure 2 Label precision for each stimulus and both groups. Each vertical bar is the cumulative percentage of precise noun, odor family and vague descriptor. Stimuli are ranked in ascending order according to the average of the familiarity scores of the experienced panel.

Previous works have shown that this odor recognition performance can be improved by cognitive factors such as familiarity with the odor. Such an effect was observed on our data, whichever the group. Unknown odors were recognized less than the more familiar ones. However, well known odorants did not lead to better recognition performances, conversely to the Rabin and Cain study (1984). Lawless and Cain (1975) noticed no influence of familiarity on d' scores. Rabin and Cain (1984) explained these results partly by the nature of the stimuli used by Lawless and Cain, i.e. chemical compounds, but also because subject variability was not taken into account in their computations of performances. Some unfamiliar chemical odors were voluntarily chosen in our experiment in order to test experienced subjects on their potential ability to memorize unfamiliar odors; moreover, d' was computed as Rabin and Cain recommended in their paper. The fact that the experienced panel performed poorly with unknown odorants in our study disproved our hypothesis that they should demonstrate better analytical abilities to encode odors and thus perform better at the recognition task, even for unknown odors. Some chemical compounds had an unusual and also complex odor, which could not be defined by only one word. People had no prior name–odor association available, nor the ability to create it quickly in these conditions (as explained by Cain, 1979), even after the sensory profiling training. Moreover, some confusion between odor quality of targets or between targets and distractors could occur and disturb subjects in their evaluation. At the end of the second session, people from both groups complained about the difficulty of the task because some odors appeared similar. After checking, we found minor confusions between tarragon and anethol, camphor and menthol. Although many authors have underlined the importance of using distinguishable and identifiable stimuli in such experiments (see Cain, 1979, 1982), no similarity measurements had been made prior to this experiment. Stimuli were smelled by the experimenter, who did not mentioned a high risk of confusion. Nevertheless, d' scores are comparable to scores obtained by

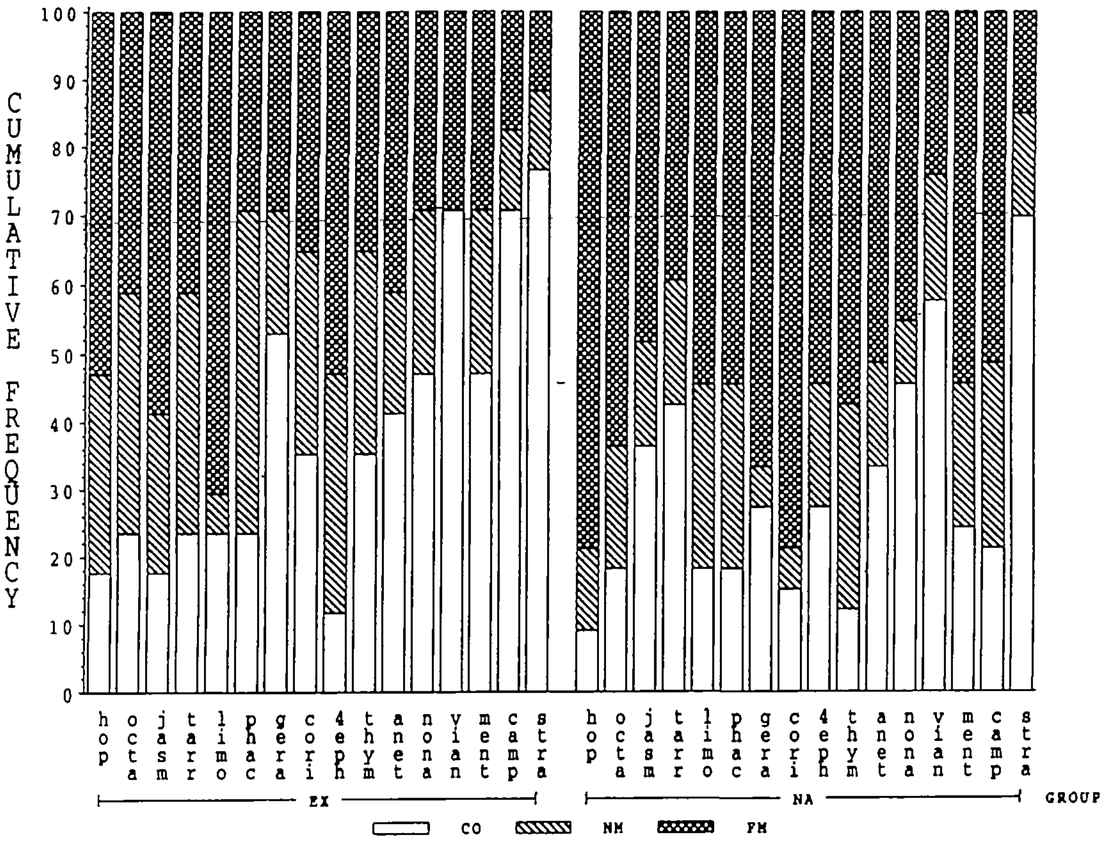

Figure 3 Label use consistency for each stimulus and both groups. Each vertical bar is the cumulative percentage of consistent, near miss and far miss label. Stimuli are ranked in ascending order according to the average of the familiarity scores of the experienced panel.

authors who used stimuli previously verified for their discriminability (Lyman and McDaniel, 1986). Finally, the choice of stimuli was rather difficult because experienced subjects should not have been exposed previously to the tested odors. These elements could mask the importance of familiarity for the recognition task but another effect was noticed on the quality of the answer at the encoding task. Indeed, some effects of panel experience were found on this answer quality. Lyman and McDaniel (1986) showed the importance of the nature of the encoding task on the strength of odor memory and demonstrated that giving a consistent label plus a definition to the perceived odor improved memory performances. In the present experiment, only labeling was required because it was a common exercise practised by the experienced panel; it was thus included as a potential influencing factor, as well as the precision of the label given at the inspection phase. A positive effect of descriptive training was effectively observed: experienced panelists gave more consistent labels (in percentage) than did naives, and they tended to generate more precise labels. According to Rabin and Cain (1984), we would expect better recognition performances from the experienced panelists. Although Engen (1987) argued that the associative power of an odor with a name is particularly weak, we observed that the profiling training program had improved subjects' ability to analyze a perception and to verbalize it more precisely. One direct consequence of this kind of training is to improve the degree of accuracy of verbal description, which is clearly linked to accuracy of perception. This is in accordance with previous studies which have shown that the accuracy of perception was strongly associated to label use consistency. However, these greater abilities did not help experienced panelists to perform better at the recognition task. The encoding process of some stimuli was probably not stored deeply enough to be correctly retrieved at the recognition phase, 7 days later. According to Davis (1975), the encoding process is more efficient when odorants are very familiar, which is due to pre-existing cognitive systems. Despite this, no direct consequence was observed on the global memory performances (d'). The two abilities of precision and consistency did not appear totally related. Indeed, the percentage of PN descriptors used consistently was no greater for the experienced panel than for the naives (see Table 6). Consequently, the impact of these two abilities on the experienced panel recognition performance was not significant, because their effects on d' were not additive. These results suggest that the verbal encoding is not the major part of the process involved in olfactory memory, as was stated by Lyman and McDaniel (1990) and Murphy *et al.* (1991). More recently, Cain and Potts (1996) produced some evidences to the coexistence of a perceptual and semantic encoding process, showing that 'errors of identification occur at input into memory as well as at the output from memory'.

Conclusion

The strength of odor memory is basically related to the previous experience of the subject with that odor; the more deeply the information is stored, the more easily it is retrieved in a recognition task (Engen, 1987). The influence of panel experience is weak, i.e. people did not learn to create a strong association in a short time to memorize odors in laboratory conditions. However, the answer quality of the encoding task is of primary importance. The semantic processing is an effective aid to store olfactive information (Murphy *et al.*, 1991), but is not sufficient. Sensory training enhances the ability of subjects to verbalize odorant stimuli accurately. However, it is not enough to enhance odor recognition performance in such conditions.

This experiment also led us to consider the data from the point of view of a descriptive panel leader. Some naive subjects obtained d' values similar to those of experienced panelists. One of the aims of the panel leader was to train selected panelists to reach a high level of performance on profiling measurements. This is partly related to a great ability to memorize odor descriptors. It is suggested that selecting people who obtained high scores at an odor recognition task could contribute to the attainment of the expected good performances. The predictive value of d' for future performances in a descriptive profiling program is unknown, but some preliminary results tend to support this idea (Lesschaeve and Issanchou, 1995).

ACKNOWLEDGEMENTS

We thank Dr Martha Petro-Turza (CFRI, Hungary), Firmenich SA (Switzerland), Ets. Robertet (France) and Kalsec (USA) for providing some of the tested odorants. Catherine Juffard, Pierre Lavier and Philippe Grégoire are thanked for their technical assistance. We thank the referee for the fruitful comments.

REFERENCES

AFNOR (1987) Analyse sensorielle. *Directives générales pour l'implantation de locaux destinés à l'analyse sensorielle.* AFNOR, Paris, France.

Amerine, M.A., Pangborn, R.M. and Roessler, E.B. (1965) *Principles of Sensory Evaluation of Food.* Academic Press, New York.

ASTM (1981) *Guidelines for the Selection and Training of Sensory Panel Members.* ASTM Committee E-18 STP 758, Philadelphia, PA.

Banks, W.P. (1970) Signal detection theory and human memory. *Psychol. Bull.*, **2**, 81–99.

Cain, W.S. (1979) To know with the nose: keys to odor identification. *Science*, **2**, 467–470.

Cain, W.S. (1982) Odor identification by males and females: predictions vs performance. *Chem. Senses*, **2**, 129–142.

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.

Davis, R.G. (1975) Acquisition of verbal associations to olfactory stimuli of varying familiarity and to abstract visual stimuli. *J. Exp. Psychol.: Hum. Learn. Mem.*, **2**, 134–142.

Desor, J.A. and Beauchamp, G.K. (1974) The human capacity to transmit olfactory information. *Percept. Psychophys.*, **3**, 551–556.

Engen, T. (1971) Olfactory psychophysics. In Beidler, L.M. (ed.), *Handbook of Sensory Physiology—Chemical Senses.* Springer-Verlag, New York, pp. 216–244.

Engen, T. (1987) Remembering odors and their names. *Am. Scient.*, **75**, 497–503.

Engen, T. and Ross, B.M. (1973) Long-term memory of odors with and without verbal descriptions. *J. Exp. Psychol.*, **2**, 221–227.

ISO (1993) Analyse sensorielle. In *Guide général pour la sélection, l'entraînement et le contrôle des sujets. Partie 1: sujets qualifiés.* AFNOR, Paris.

Issanchou, S., Lesschaeve, I. and Köster, E.P. (1995) Screening individual ability to perform descriptive analysis of food products. Basic statements and application to a Camembert cheese descriptive panel. *J. Sens. Stud.*, **10**, 349–361.

Lawless, H.T. and Cain, W.S. (1975) Recognition memory for odors. *Chem. Senses*, **1**, 331–337.

Lesschaeve, I. and Issanchou S. (1995) Could selection tests detect the future performances of descriptive panelists? *Second Pangborn Sensory Symposium*, 30 July–3 August, Davis, CA.

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

Lyman, B.J. and McDaniel, M.A. (1990) Memory for odors and odor names: modalities of elaboration and imagery. *J. Exp Psychol.: Learn. Mem. Cogn.*, **4**, 656–664.

Meilgaard, M., Civille, G.V. and Carr, B.T. (1991) *Sensory Evaluation Techniques.* CRC Press, Boca Raton, FL.

Murphy, C., Cain, W.S., Gilmore, M.M. and Skinner, R.B. (1991) Sensory and semantic factors for recognition memory for odors and graphic stimuli: elderly versus young persons. *Am. J. Psychol.*, **2**, 161–192.

Rabin, M.D. (1988) Experience facilitates olfactory quality discrimination. *Percept. Psychophys.*, **6**, 532–540.

Rabin, M.D. and Cain, W.S. (1984) Odor recognition: familiarity, identifiability and encoding consistency. *J. Exp. Psychol.: Learn. Mem. Cogn.*, **2**, 316–325.

SAS (1988a) *SAS/STAT—User Guide*. SAS Institute, Cary, NC.

SAS (1988b) *SAS/GRAPH—User Guide*. SAS Institute, Cary, NC.

Stone, H. and Sidel, J.L. (1985) Food science and technology. In *Sensory Evaluation Practices.* Academic Press, London.

Stone, H., Sidel, J., Oliver, S., Woolsey, A. and Singleton, R.C. (1974) Sensory evaluation by quantitative descriptive analysis. *Food Technol.*, Nov., 24–34.

Received on September 12, 1994; accepted on July 24, 1996