Chem. Percept.
DOI 10.1007/s12078-009-9050-8

Beer-Trained and Untrained Assessors Rely More on Vision than on Taste When They Categorize Beers

Maud Lelièvre · Sylvie Chollet · Hervé Abdi · Dominique Valentin

Received: 20 April 2009 / Accepted: 9 July 2009

Abstract What role categorization processes play in chemosensory expertise and its acquisition? In this paper, we address this question by exploring the criteria used by trained and untrained assessors when they categorize beers. Two experimental factors were manipulated: beer color and brewery. Participants sorted nine commercial beers coming in three different colors and from three different breweries. Participants sorted in two different conditions: in one condition, participants could see the beers, and in the other condition, they could not see the beers. We observed that in both tasting conditions (i.e., with or without vision), trained and untrained assessors categorized beers similarly. In the visual condition, assessors sorted beers by color, whereas in the blind condition, they sorted them by brewery. Overall, our results indicate that sensory training does not seem to have an effect on the criteria used to organize beer perceptions. This suggests that our trained beer assessors did not develop specific conceptual representations of beers during training. Moreover, it seems that when assessors categorize beers, they rely more on visual than on chemosensory information.

M. Lelièvre (✉) · S. Chollet
Institut Supérieur d'Agriculture,
48 Boulevard Vauban,
59046 Lille Cedex, France
e-mail: m.desmas@isa-lille.fr

M. Lelièvre · D. Valentin
UMR CSG 5170 CNRS, Inra, Université de Bourgogne,
21000 Dijon, France

H. Abdi
The University of Texas at Dallas,
Richardson, TX 75080-3021, USA

Keywords Beer · Categorization · Expertise · Training · Visual Information

Introduction

Perfumers, oenologists, and beer experts possess an expertise that set them apart from other people, even perfume, wine, or beer lovers or aficionados. This professional expertise relies both on chemosensory and a technical knowledge. The chemosensory knowledge is acquired through repeated tastings and sniffings of beer, wine, or perfumes. The technical knowledge encompasses knowledge in chemistry, viticulture, oenology, brewing processes, etc. Understanding expertise relative to chemical senses is a rather new field of research compared to understanding expertise in other fields such as physics, problem solving, computer programming, games, sport, or medicine (for reviews, see Vicente and Wang 1998; Feldon 2007). In the sensory field, the notion of expertise is also used, but with another definition. Beer or wine experts in sensory evaluation are assessors trained to evaluate the intensity of different attributes of the products and to detect and identify flavors and defects. Their expertise relies mainly on sensory expertise. For clarity, in this paper, professional experts will be called "experts" and experts in sensory evaluation will be called "trained assessors."

Chemosensory expertise has been relatively well studied in terms of discrimination, description, and, to a lesser extent, memory performance (see Chollet and Valentin 2000, 2006; Labbé et al. 2004; Valentin et al. 2007 for reviews). Overall, it seems that both experts and trained assessors are better than novices to describe, memorize, and discriminate between stimuli, but the difference in performance is not always very impressive.

Published online: 23 July 2009

In addition, a small number of studies on wine expertise dealt with the organization of experts' knowledge and its influence on chemosensory perception (Solomon 1997; Hughson and Boakes 2002; Ballester et al. 2005, 2008). A first study by Solomon (1997) suggested that when we acquire wine expertise, we create a system of categorization based on the appreciation of salient characteristics of wines. In this study, experts, intermediate, and novice wine assessors were asked to sort ten white wines from different grape types into four groups by putting together the wines they perceived similar. The results revealed that experts tended to categorize the wines by grape types. By contrast, intermediate and novice wine assessors did not sort the wines by grape types but rather used one or two salient perceptual characteristics to explain their sort. More recently, Ballester et al. (2008), using a typicality rating task, showed that for wine experts, Chardonnay and Muscadet wines were organized along a typicality gradient going from very typical of a Chardonnay (or a Muscadet) wine to very atypical of this wine. No such gradient was observed for novices. The authors concluded that contrary to novices, experts have common separate mental representations of Chardonnay and Muscadet wines and that these representations are partially based on perceptual similarities. These studies suggest that due to their knowledge on wine and through repeated wine tastings, wine experts—contrary to novices—develop specific and consensual mental representations of wines.

No such studies have been run with trained assessors. Yet trained assessors, although they do not possess the vast technical knowledge of experts, are routinely used to taste food products. A typical tasting session consists in a formal training where assessors learn to recognize and describe perceptual characteristics in a repeatable and consistent manner. Moreover, assessors generally receive some information on the products they taste. So it is possible that through these repeated exposures, trained assessors extract sensory characteristics common to different products and develop concepts different from those of novices. So we could expect that a modification of mental representations would also occur with sensory training. Chollet and Valentin (2001, 2006) failed to find an effect of training on beer categorization, but this result could be due to the short training (11 and 15 h, respectively) of their assessors. To test this possibility, we evaluated whether beer assessors trained for several years develop specific and consensual beer sensory concepts compared to novices.

An examination of beer tasting manuals (Jackson 1999; d'Eer 1998, 2005; Glover 2001; Mastrojanni 1999; Webb 2005) shows that there are many criteria for categorizing beers. The most usual and easiest categorization scheme is based on color: blond, amber, dark, white, and red beers. Other usual categorization schemes are based on country of origin (e.g., Belgian, French, English, and German beers) or alcohol content—but this system of classification differs from country to country. We also find categorization schemes that are less accessible to beer consumers because these categorizations are based on technical criteria, such as fermentation type, raw materials, packaging methods, brewer characteristics, etc.

In the present study, we explored the effect of training on two different categorization criteria, beer color and brewery, and tested these criteria in two different tasting conditions, with and without visual information. We hypothesized that untrained assessors would categorize beers only according to perceptual similarities, whereas beer-trained assessors would rely more heavily on their knowledge of beer sensory characteristics. Because we assumed that categorizing beers by colors required less knowledge than categorizing them by brewery, we expected that untrained assessors would sort beers by color and that trained assessors would sort by brewery. Moreover, as trained assessors are used to evaluate beer under red light and to focus their attention on beer odor and flavor, we expected that in the visual condition, trained assessors would be less influenced by beer color than untrained assessors would. As a result, we hypothesize that the difference between trained and untrained assessors would be larger in the visual than in the blind condition.

Finally, because there is clear individual variability in categorization tasks (Chollet and Valentin 2001; Ishii et al. 1997), the stability of the categorization data in the blind condition was evaluated in two ways. First, we examined the stability within individuals: we analyzed the repeatability of beer categorizations of each assessor over four repetitions of the sorting task. Second, we examined the stability between individuals to check the agreement among participants by (1) comparing beer categorizations of three groups of assessors (two groups of untrained and one group of trained assessors) and (2) studying the consensus within the trained and untrained groups for the first repetition.

Material and Methods

Assessors

In the visual condition, two groups of assessors participated to the task: one group of 17 trained assessors and one group of 21 untrained assessors. In the blind condition, three groups of assessors participated to the task: one group of 13 trained assessors and two groups of 18 (group A) and 37 (group B) untrained assessors.

The trained assessors were part of a sensory beer panel and were staff members from the Catholic University of Lille (France). This panel was slightly modified between the experiments in the visual and blind conditions as some assessors integrated the panel and others left it. This explained the different number of trained assessors between the two conditions. However, 11 of the trained assessors were common to the two conditions. Trained assessors had been formally trained to evaluate different kind of beers (including the beers studied here) 1 h per week for an average of 3.5 years. They were trained to detect and identify flavors (almond, banana, butter, caramel, cabbage, cheese, lilac, metallic, honey, bread, cardboard, phenol, apple, and sulfite) added in beer and to evaluate the intensity of general compounds (bitterness, astringency, sweetness, alcohol, hop, malt, fruity, floral, spicy, sparklingness, and lingering). The untrained assessors were students or staff members from the Catholic University of Lille and the University of Bourgogne (France). Untrained assessors were beer consumers who did not have any formal training or experience in the description of beers. Group A untrained assessors included 18 assessors in order to be compared to the trained assessors group. Group B untrained assessors was larger in order to evaluate the stability between individuals.

The number of assessors, the proportion of women/men, the mean age, and the mean training time for trained and untrained assessors in each condition are presented in Table 1.

Products

Nine different commercial beers were evaluated (denoted PelfBL, PelfA, PelfBR, ChtiBL, ChtiA, ChtiBR, LeffBL, LeffA, and LeffBR). These beers came from three different breweries, *Pelforth* (noted Pelf), *Chti* (Chti), and *Leffe* (Leff), and each brewery provided three types of beer: blond (BL), amber (A), and dark (BR). All beers were presented in three-digit coded black plastic tumblers and served at 10 °C.

Procedure

Visual Condition

The experiment was conducted in separate booths lighted with a neon lighting of 18 W. Assessors took part individually in the experiment which was run in a single session. They were provided with the nine beers and were required to sort them by putting together the beers they perceived similar. They were allowed to smell and taste samples many times. No criterion was provided to perform the sorting task. Assessors were free to make as many groups as they wanted and to put as many beers as they wanted in each group. They were allowed to take as much time as they wanted. Mineral water and bread were available for assessors to rinse between samples. Assessors could spit out beers if they wanted.

Blind Condition

The experiment in the blind condition took place 1.5 years before the experiment in the visual condition. The experiment was conducted in separate booths lighted with a neon lighting of 18 W with a red filter darkened with black tissue paper to mask the color differences between beers. Group B untrained assessors took part individually in a single session consisting in a sorting task on the nine beers as described in the previous section. Trained assessors and group A untrained assessors carried out four repetitions of this sorting task. Repetitions 1 and 2 were conducted during a first session and were separated by a 20-min break. Repetitions 3 and 4 were conducted during a second session which took place 1 week later.

Data Analysis

Beer Similarity Spaces

For each assessor, the results of the sorting tasks were encoded in individual distance matrices where the rows and the columns are beers and where a value of 0 between a row and a column indicates that the assessor put the beers together, whereas a value of 1 indicates that the beers were not put together. For each group of assessors and each condition (visual and blind conditions), the individual distance matrices obtained from the sorting data were analyzed with DISTATIS (Abdi et al. 2005, 2007). DISTATIS is a three-way generalization of classical multidimensional scaling which takes into account individual sorting data. This method provides, for the products, a multidimensional scaling-like compromise map which is obtained from a principal component analysis performed on the DISTATIS compromise cross-product matrix. This compromise matrix is a weighted average of the cross-product matrices associated with the individual distance matrices derived from the sorting data (Abdi et al. 2007). In this map, the proximity between two points reflects their similarity. In addition, hierarchical ascending classifications (HAC) with Ward's criterion were performed on product coordinates to identify groups of beers on each configuration.

Comparison of Two Configurations

To compare two configurations, we computed R_V coefficients between product coordinates of each configuration. The R_V

Table 1 Composition of the assessors' groups for each condition

	Visual condition		Blind condition		
	Trained assessors	Untrained assessors	Trained assessors	Untrained assessors Group A	Untrained assessors Group B
No. of assessors	17	21	13	18	37
	7 women/10 men	7 women/14 men	5 women/7 men	5 women/13 men	15 women /22 men
Mean age (years)	38.5	36.8	34.6	29.3	25.6
Age range (years)	26–56	22–60	24–54	20–70	21–56
Mean training (years)	3.6		3.4		
Range of training (years)	1–6.5		2–5		

coefficient measures the similarity between two configurations and can be interpreted in a manner analogous to a squared correlation coefficient (Escouffier 1973), but its test requires a specific procedure (see Abdi 2007 for details).

Results

DISTATIS Similarity Spaces in the Visual Condition

Table 2 presents the eigenvalue and the percentages of explained variance of each of the nine dimensions for trained and untrained assessors. For both groups of participants, four dimensions were selected as the most appropriate solution (71.4% and 70.4% of the total variance for trained and untrained assessors, respectively). Figure 1 shows the four-dimension compromise maps obtained for trained and untrained assessors. Ellipsoids identify the product clusters obtained with a hierarchical ascending classification on the coordinates of the products for subspaces 1–2 and subspaces 3–4 for both trained and untrained assessors (Fig. 2). For both groups of assessors, the first dimension tends to oppose the three dark beers (PelfBR, LeffBR, and ChtiBR) and the LeffA to the Chti and Pelforth blond beers. The second dimension tends to oppose the three amber beers (PelfA, LeffA, and ChtiA) to the other beers. Globally, the HAC (Fig. 2) reveals a categorization by beer color—blond, amber, and dark—with two exceptions: trained assessors categorized the LeffBL with the three amber beers, whereas untrained assessors categorized the LeffA with the three dark beers. This similarity between the first subspaces (left of Fig. 1) of trained and untrained assessors is confirmed by the large value of the R_V coefficient computed between these two configurations (R_V=0.70, p=0.0007). The second subspaces (right of Fig. 1) are different for trained and untrained assessors (R_V=0.22, p=0.80), and for both assessors' groups, no categorization criteria (brewery or color) appear when using the results of the HAC.

DISTATIS Similarity Spaces in the Blind Condition

Table 3 presents the eigenvalue and the percentage of explained variance of each of the nine dimensions for trained and untrained assessors. For both groups of participants, four dimensions were selected as the most appropriate solution (69.9% and 65.8% of the total variance for trained and untrained assessors, respectively). Figure 3 shows the four-dimension compromise maps for products obtained for trained and untrained assessors in the first repetition of the sorting task. In each map, ellipsoids identify the product clusters obtained with HAC (Fig. 4).

Overall, similar clusters were observed on the first two dimensions for trained and untrained assessors. This observation is confirmed by the significant value of the R_V coefficient (R_V=0.58, p=0.02). In the first subspaces (left of Fig. 3), beers tend to be categorized by breweries, suggesting that there are more perceptual similarities between beers from the same brewery than between beers from different breweries. For trained assessors, the three Leffe beers are in the same group and the Pelforth beers and the Chti beers tend to be grouped together, respectively, with an inversion for the ChtiBR and the PelfBL. For untrained assessors, the three Chti beers are in the same group, and the Pelforth beers and the Leffe tend to be grouped together, respectively, with two exceptions: PelfBL is grouped with the Chti beers and LeffBL is grouped with PelfBR and PelfA. The second subspaces (right of Fig. 3) are different for trained and untrained assessors (R_V=0.11, p=0.28). For trained assessors, dimension 3 opposes dark beers to amber beers with blond beers in the middle. This pattern suggests that trained assessors tended to categorize beers by color. However, this observation should be interpreted with caution since dimension 3 explains only a relatively small

Table 2 Eigenvalues, percentages of explained variance, and cumulated variance of the nine dimensions for trained and untrained assessors for the sorting task in the visual condition

	Trained assessors			Untrained assessors		
	Eigenvalues	Variance (%)	Cumulated variance (%)	Eigenvalues	Variance (%)	Cumulated variance (%)
Dimension 1	0.66	26.78	26.78	0.59	26.36	26.36
Dimension 2	0.43	17.58	44.35	0.42	18.98	45.34
Dimension 3	0.37	14.93	59.29	0.31	13.88	59.22
Dimension 4	0.30	12.16	71.44	0.25	11.14	70.37
Dimension 5	0.26	10.40	81.84	0.20	9.13	79.50
Dimension 6	0.22	9.09	90.93	0.19	8.67	88.17
Dimension 7	0.14	5.72	96.65	0.16	7.04	95.21
Dimension 8	0.08	3.35	100.00	0.11	4.79	100.00
Dimension 9	0.00	0.00	100.00	0.00	0.00	100.00

amount of total variance (14.7%). For untrained assessors, we do not observe such a separation by beer colors.

Comparison Between DISTATIS Spaces in the Visual and Blind Conditions

To evaluate the effect of visual information on categorization, we computed R_V coefficients between the configurations obtained in the two conditions for trained and untrained assessors, respectively. These R_V coefficients were significant neither for the trained [R_V (Dim1–2)=0.38, p=0.33; R_V (Dim3–4)=0.11, p=0.27] nor for the untrained assessors [R_V (Dim1–2)=0.09, p=0.27; R_V (Dim3–4)=0.28, p=0.82], suggesting that assessors changed their perception of the similarity between beers when they have access to visual information.

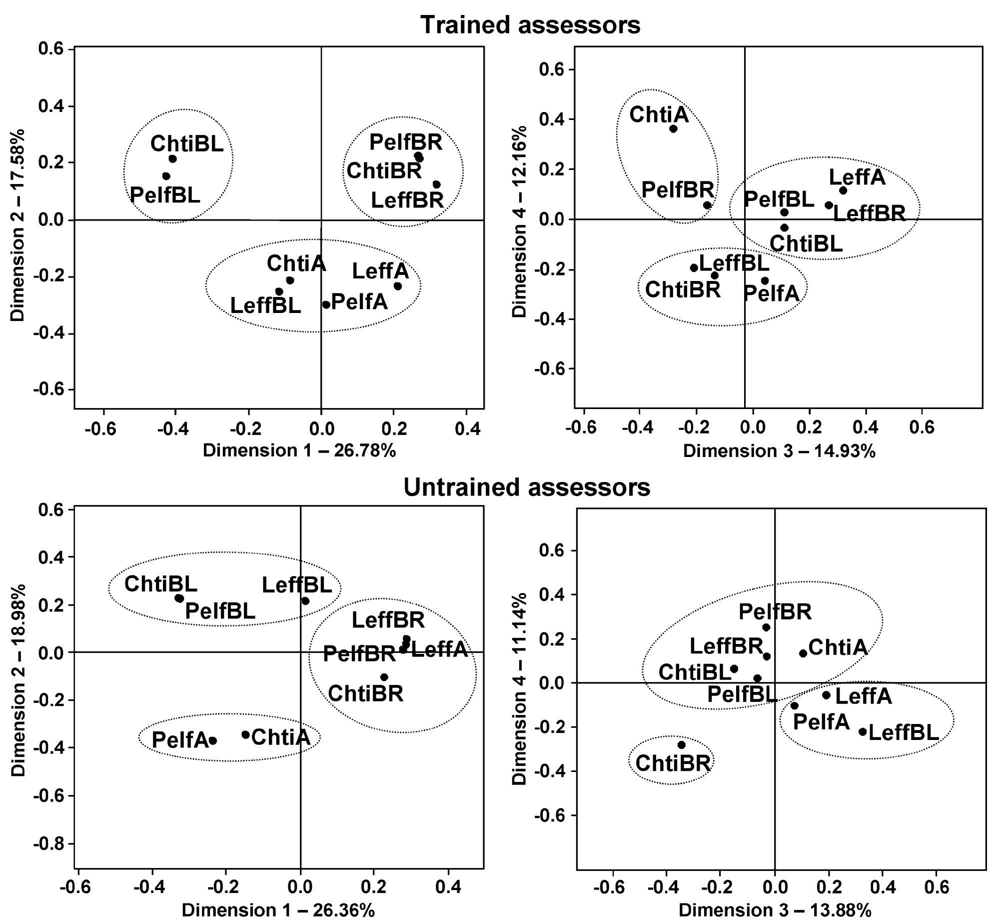

Fig. 1 Four-dimensional compromise maps for trained (*top*) and untrained assessors (*bottom*) for the sorting task in the visual condition. The *ellipsoids* correspond to the clusters identified with a HAC

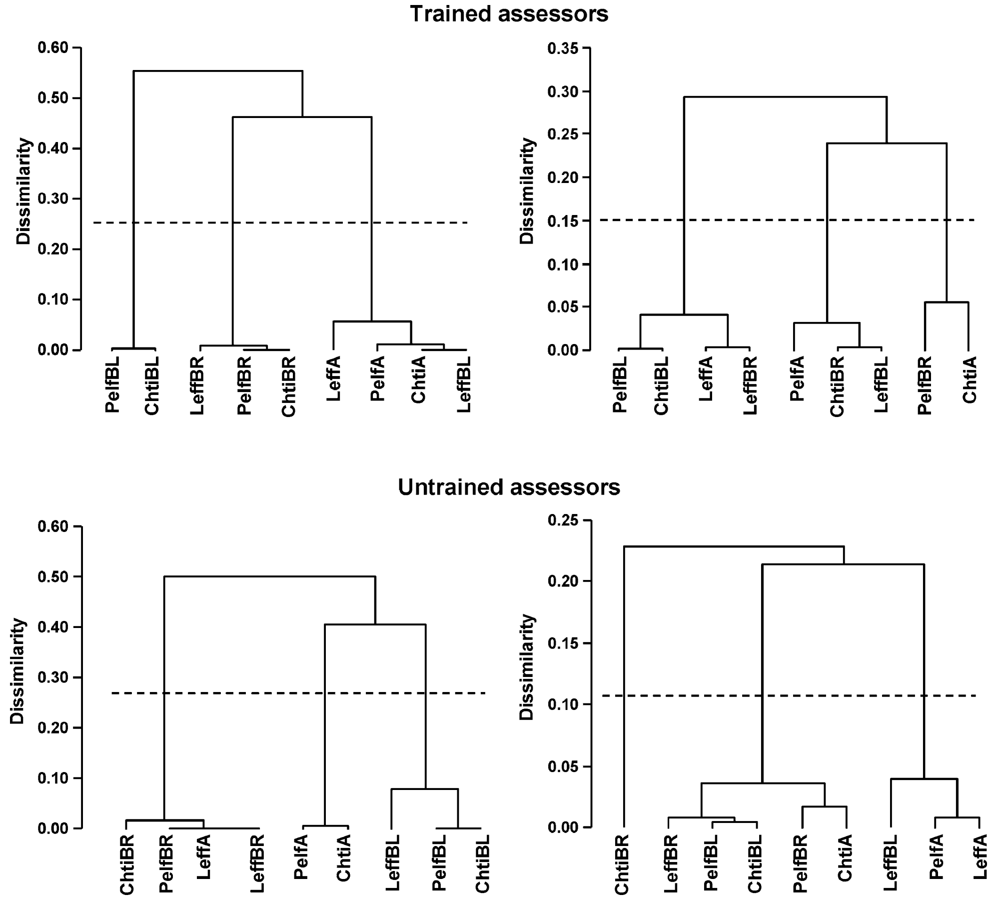

Fig. 2 Dendrograms of the HAC computed on the products' coordinates on dimensions 1 and 2 (*left*) and on dimensions 3 and 4 (*right*) for trained (*top*) and untrained (*bottom*) assessors' sorting data in the visual condition

Stability of the Sorting Data

Stability Within Individuals

We evaluated trained and group A untrained assessors' repeatability by comparing their sorts over the four repetitions of the sorting task. A two-factor analysis of variance (ANOVA) on the within-assessors R_V coefficients (i.e., the R_V coefficients were computed between each individual matrix in each repetition) was carried out to evaluate the effect of training (trained/untrained) and time between two repetitions (20 min: means of the two R_V coefficients computed between repetitions 1–2 and 3–4/ 1 week: means of the four R_V coefficients computed between repetitions 1–3, 1–4, 2–3, and 2–4). Training was entered in the model as a between-subject variable and time between two repetitions as a within-subject variable. No significant effect of training or time was found. However,

Table 3 Eigenvalues, percentages of explained variance, and cumulated variance of the nine dimensions for trained and untrained assessors for their first sorting task in the blind condition

	Trained assessors			Untrained assessors		
	Eigenvalues	Variance (%)	Cumulated variance (%)	Eigenvalues	Variance (%)	Cumulated variance (%)
Dimension 1	0.58	24.92	24.92	0.54	24.95	24.95
Dimension 2	0.44	18.69	43.62	0.37	17.06	42.00
Dimension 3	0.34	14.70	58.31	0.28	12.97	54.97
Dimension 4	0.27	11.57	69.88	0.24	10.81	65.79
Dimension 5	0.21	9.04	78.92	0.22	10.28	76.07
Dimension 6	0.21	8.88	87.80	0.20	9.11	85.17
Dimension 7	0.16	6.94	94.74	0.19	8.51	93.68
Dimension 8	0.12	5.26	100.00	0.14	6.32	100.00
Dimension 9	0.00	0.00	100.00	0.00	0.00	100.00

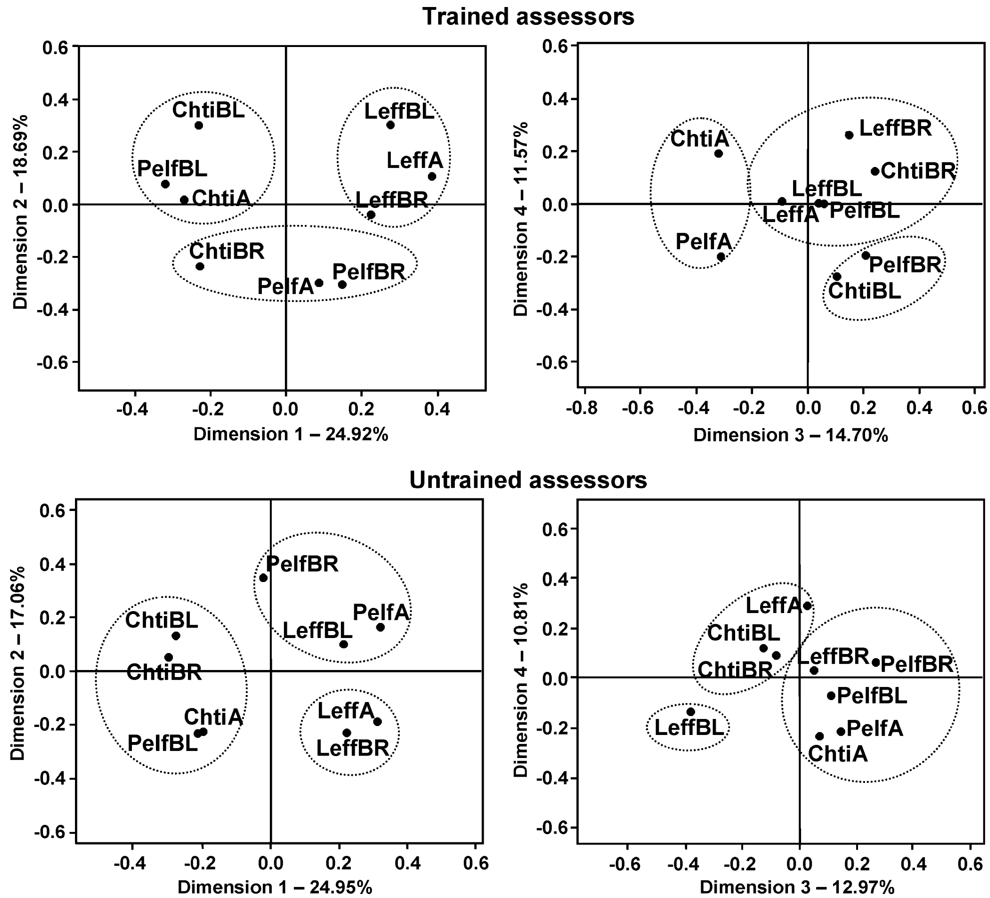

Fig. 3 Four-dimensional compromise maps for trained (*top*) and untrained assessors (*bottom*) for their first sorting task in the blind condition. The *ellipsoids* correspond to the clusters identified with HAC

this failure to show any significant effect can be explained by the low power of the test due to the relatively small number of participants and by the large inter-individual variability of the R_V coefficients. Figure 5 shows the box plots of R_V coefficients distribution calculated between each repetition for trained and untrained assessors. In this figure, we see that the inter-individual variability of the R_V coefficients for trained assessors is globally smaller than that for untrained assessors, whereas the R_V coefficient means are very similar between the two groups of assessors. To analyze further this inter-individual variability, we computed the 99% confidence interval for each R_V coefficient and counted the number of trained and untrained assessors whose R_V coefficient differed significantly from zero. We found that on all the repetitions, 94.9% of trained assessors had an R_V coefficient significantly different from zero versus 70.4% for untrained assessors. Thus, trained assessors were globally more repeatable than untrained assessors on the four repetitions even though the ANOVA failed to show any effect on the average R_V coefficients.

Stability Between Individuals

To have an estimation of the participants' agreement on their categorization in the blind condition, we first compiled the sorting data of the first repetition of the 13 trained and 18 untrained assessors of group A with the sorting data of the 37 group B untrained assessors to obtained 68 different sorts. We used a bootstrap procedure on these data to examine the stability of the categorizations over these 68 sorts. The bootstrap (Efron and Tibshirani 1993) is a non-parametric statistical inference method that is used here to create confidence intervals for the position of products. For DISTATIS, the bootstrap first creates a set of distance matrices by sampling with replacement from the initial sorting data matrices, and DISTATIS is used to compute a compromise matrix from this set of distance matrices. Then, this procedure is repeated 1,000 times and all the 1,000 compromise matrices are projected on the original compromise. A confidence ellipsoid that comprised 95% of the projections of these compromise matrices is computed for each product (see Abdi et al. 2009 for more details on these computations).

Figure 6 shows the four-dimensional compromise map. The confidence ellipsoids represent the variability of the results over the 68 assessors. We observe that ellipsoids are small, which shows that the assessors' categorizations are stable. In addition, we can see on the first two-dimension maps (Fig. 6 on the left) that the confidence ellipsoids of the three Leffe beers, of the three Chti beers, and of the

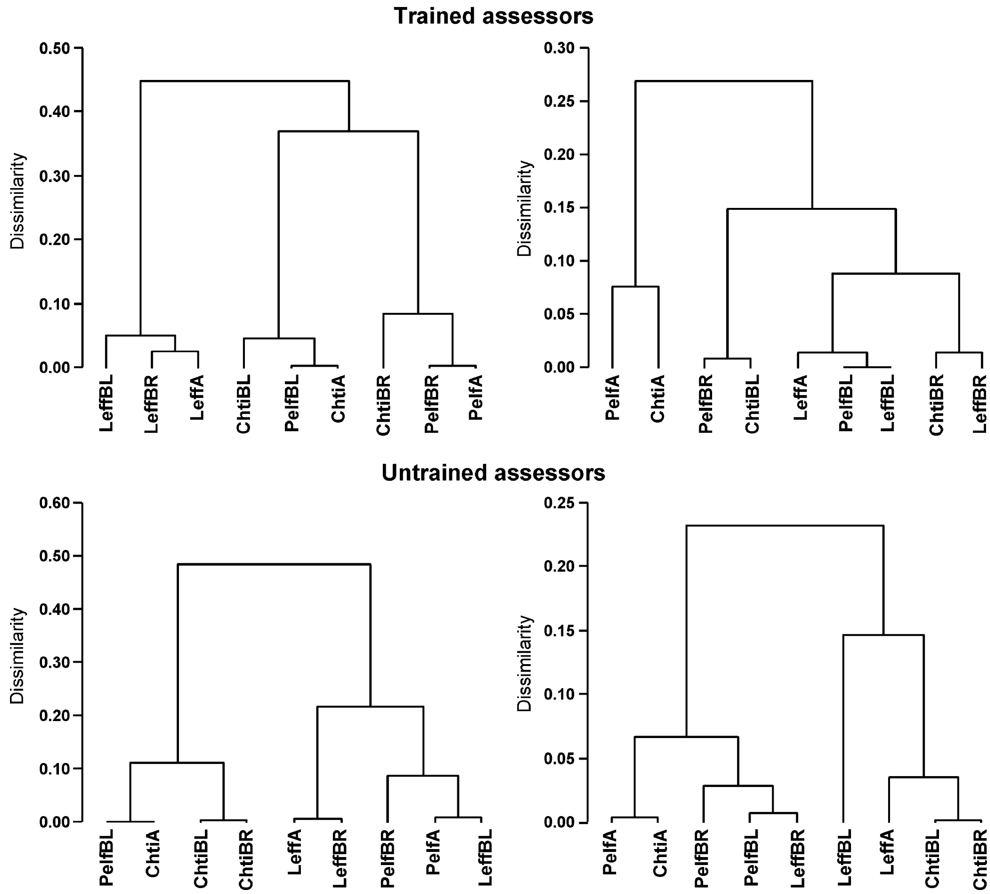

Fig. 4 Dendrograms of the HAC computed on the products' coordinates on dimensions 1 and 2 (*left*) and on dimensions 3 and 4 (*right*) for trained (*top*) and untrained (*bottom*) assessors' sorting data in the blind condition

PelfBR and PelfA overlap, respectively, but do not overlap between them. Such a pattern confirms that trained and untrained assessors globally categorized beers mostly by using the brewery criterion. On dimensions 3 and 4 (Fig. 6 on the right), confidence ellipsoids of the nine beers overlap, except those of PelfA and PelfBL. This pattern confirms that assessors are not totally in agreement on the categorizations. Yet, we observe that beers tended to be categorized by beer colors on these dimensions 3 and 4. In fact, the confidence ellipsoids of the three dark beers are very close and are opposed to the confidence ellipsoids of blond and amber beers.

Second, to evaluate the effect of training on between-individual stability, we examined the consensus within trained and group A untrained assessors by computing R_V coefficients between the individual matrices of each assessor and the rest of her/his group for the first repetition (Fig. 7). A Student's t test performed on these R_V coefficients showed a significant difference in consensus between trained and untrained assessors [$t(29)=2.18$; $p<$

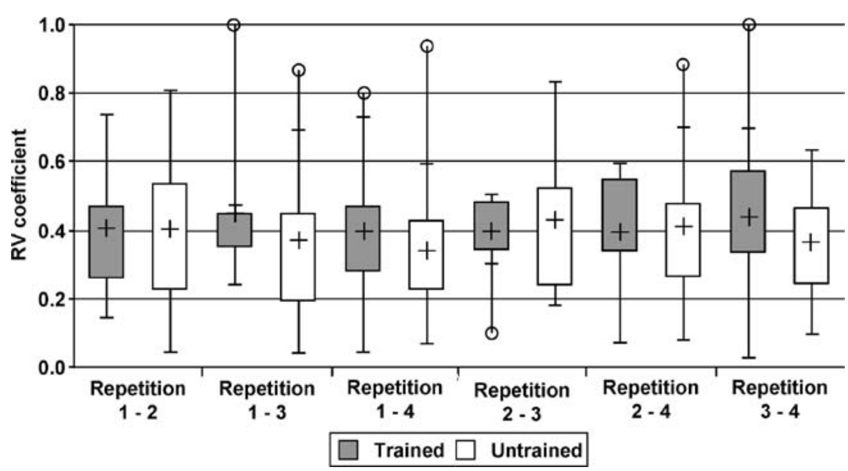

Fig. 5 Box plot of R_V coefficients distribution calculated between each of the four repetitions for trained (*black boxes*) and untrained (*white boxes*) assessors. The *box* extends from the first to the third quartile, the *plus sign* represents the mean value, and the *ends of the lines extending from the box* (*whiskers*) indicate the maximum and the minimum data values, unless outliers are present in which case the *whiskers* extend to a maximum of 1.5 times the inter-quartile range (i.e., length of the box)

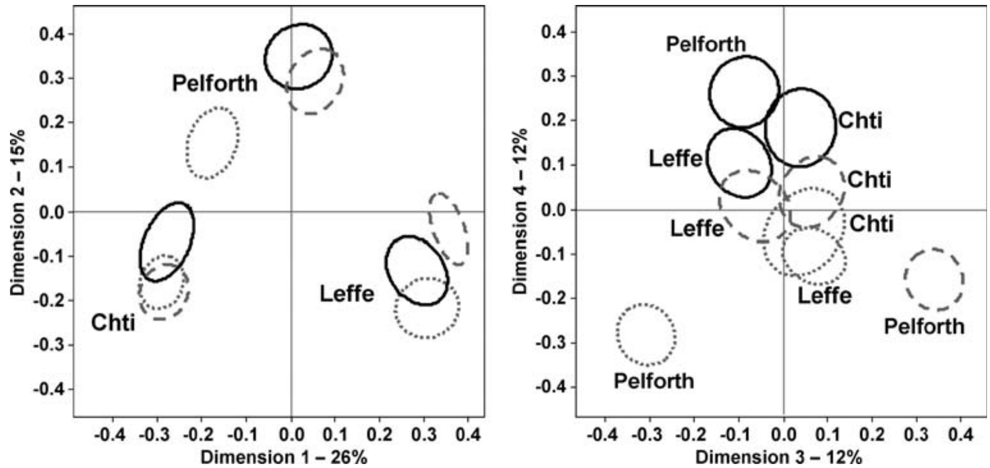

Fig. 6 Four-dimensional compromise maps of the products with 95% confidence ellipsoids computed by bootstrap on trained assessors and the two groups of untrained assessors' data. The *solid black lines* represent the dark beers, the *dashed gray lines* represent the amber beers, and the *dotted gray lines* represent the blond beers

0.05]. This shows that trained assessors were more consensual than untrained assessors on their categorization.

Discussion

With this experiment, we wanted to explore what categorization criteria are used by trained and untrained beer assessors when they organize their perceptions of beers. We used sorting tasks in order to study two different criteria, beer color and brewery, in two tasting conditions, visual and blind conditions. We expected that the organization of beer perceptual spaces would be different for trained and untrained assessors because trained assessors have knowledge about beer sensory characteristics that could affect their categorization criteria.

Influence of Visual Information on Beer Categorization

In the visual condition, trained and untrained assessors both clearly categorized beers by color, whereas in the blind condition, they categorized beers by brewery. This result suggests that in a beer categorization task, both trained and untrained assessors tend to rely more on visual information than on chemosensory information. While this visual dominance is not surprising for untrained assessors, we expected trained assessors to be less influenced by beer color because they are used to taste beer under red light and to focus their attention on beer taste and aroma. Moreover, the trained assessors declared that they had not based their sort on visual information but on beer chemosensory properties. Yet, such an influence of color on taste experts' chemosensory perception has been previously reported. For example, Pangborn et al. (1963) found an effect of color on the sweetness perception of wines: a white wine colored pink (in order to give it the appearance of a blush wine) was perceived by wine experts as sweeter than the same uncolored wine. More recently, Morot et al. (2001) showed that wine experts described the odor of a white wine artificially colored in red as a red wine. The results of the experiment of Morrot et al. demonstrate that experts rely on visual appearance for interpreting chemosensory information. In other words, humans are so visually oriented that even experts look for visual cues when they interpret chemosensory information, and these visual cues may mask other information as we observed here. This masking phenomenon can be explained either in terms of selective attention or in terms of congruence seeking. From a selective attention perspective, we can hypothesize that assessors could not process both visual and chemosensory information at the same time and that they performed the categorization task mostly on visual information, tuning out smell, and taste information. From a congruence-seeking perspective, we can hypothesize that assessors can switch their attention from visual information to chemosensory information but that chemosensory perception is driven by visual information. That is, when attending to chemosensory information, assessors will, unconsciously, seek a confirmation of visual information rather than analyzing smell and taste properties. This second explanation might

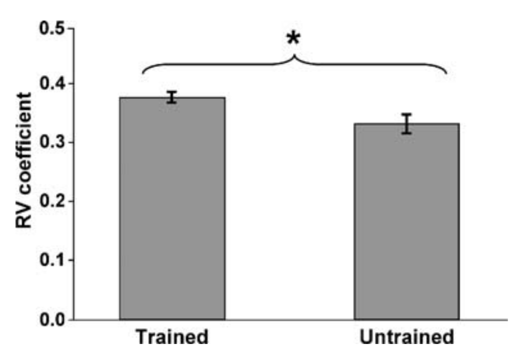

Fig. 7 Means and standard errors of R_V coefficients computed between the individual matrices of each assessor and the rest of her/his group for trained and untrained assessors for the first repetition

explain why our trained assessors felt that they relied more on chemosensory information than on visual information when they actually relied more on visual information.

Finally, the dominance of breweries in beer categorization observed in the blind condition suggests the existence of a brewery pattern which is perceived by both trained and untrained assessors. This brewery pattern seems to be common to all types of beers (blond, amber, and dark). This result is reminiscent of a study by Aubry et al. (1999) who found a winery pattern for Burgundy Pinot noir wines. The authors showed that trained assessors unconsciously grouped the wines from a given winemaker maybe because each winemaker has a specific "style." In the case of beers, this brewery pattern could be explained by several brewing parameters (brewing temperature, time, type of filtration, pasteurization, etc.), among which the yeasts could be an important factor conferring to each brewery a specific style.

Influence of Expertise on Beer Categorization

Contrary to our hypothesis, we did not observe an effect of training on beer categorization. Following Solomon (1997) and Ballester et al. (2008), we expected trained assessors to rely more on conceptual information (brewery) than on perceptual information (color). In Solomon's study, whereas experts categorized wines according to grape variety, novices used "surface characteristics" such as sweetness or fruitiness to sort the wines. The author explained this difference between experts and consumers in terms of learning effects. Through repeated exposures to wines made from different grape varieties, experts have developed wine representations based on grape variety. When asked to perform a sorting task, experts would then use these wine representations to guide their choices. Consumers do not have such representations, and therefore, they would base their sorting decisions on surface characteristics such as sweetness. In other words, according to Solomon, experts and consumers would use different cognitive processes to perform a sorting task. Experts would rely heavily on top-down information (memory representation), whereas consumers would rely more heavily on bottom-up information (perceptual or surface characteristics).

One reason to explain why we did not observe an effect of training on categorization criteria could be that—as mentioned in "Introduction"—in comparison with the studies of Solomon (1997) and Ballester et al. (2008), our trained assessors are not beer experts in the sense of being "professionals": their professions are not connected with beer. They regularly attend a formal training in sensory analysis of beer in which they learn to recognize, identify, and communicate the sensory properties of beers but they do not possess academic knowledge about beer process or marketing. This standard sensory expertise is quite different from the professional expertise of the participants of Solomon (1997) and Ballester et al. (2008) In fact, for winemakers, wine professionals, sommeliers, students in Oenology, and other wine professionals, categorization is the base of training. These wine professionals use the notions of grape type, region of origin, year, etc. to communicate about and to categorize wines, and they also are trained to recognize a specific wine as a member of a category. This could explain why the professionals of Solomon (1997) and Ballester et al. (2008) used grape type as a criterion to categorize wines. In contrast, our trained assessors did not learn a beer categorization system and so were not used to think of a beer in terms of membership to a specific category. They learned to describe the perceptual characteristics of beers in a repeatable and consensual manner. So it seems that a mere exposure to beers is not enough to build conceptual representations. Individuals need to be engaged in a training involving categorization process to develop the common mental representations that differentiate them from novices. We can hypothesize that brewers, whose expertise seems to be more similar to those of the participants of Solomon and Ballester et al., would categorize the beers differently from untrained assessors. Moreover, contrary to wine consumers who do not know the notion of grape type, beer consumers are familiarized with the breweries since they use it when buying beer for example. It can also explain why we did not observe a difference between trained and untrained assessors' categorizations.

Yet, even if beer sensory training does not lead to conceptual representations of beers, trained assessors are globally more repeatable and more consensual than untrained assessors, as has been already shown (Clapperton and Piggott 1979; Chollet and Valentin 2006). So it seems that sensory training does yield some changes in mental representations even if it is not enough to build conceptual representations different from novices. However, further work is needed to better understand these changes.

Conclusion

In this paper, beer categorization was studied as a way to better understand chemosensory expertise. We explored the change in the organization of beer perception with a sensory training by examining beer categorization criteria used by trained and untrained assessors. First, the experiments highlight the dominance of visual information on chemosensory information during a beer categorization task. Assessors, whatever their level of expertise, used visual cues as a base before analyzing chemosensory properties of beers.

Second, we did not observe an effect of sensory training on the organization of the system of categorization. Compared to previous studies with wine experts, it can suggest that there is a difference between professional expertise and sensory expertise. The technical knowledge professional experts have in addition to chemosensory knowledge could have a great influence on their perceptions. Moreover, the difference of training approach between professional experts and sensory assessors could explain that professionals develop common mental representations of products that differentiate them from novices in a categorization task.

References

Abdi H (2007) The R_v coefficient and the congruence coefficient. In: Salkind N (ed) Encyclopedia of measurement and statistics. Sage, Thousand Oaks, pp 849–853

Abdi H, Dunlop JP, Williams LJ (2009) How to compute reliability estimates and display confidence and tolerance intervals for pattern classifiers using the bootstrap and 3-way multidimensional scaling (DISTATIS). NeuroImage 45:89–95

Abdi H, Valentin D, O'Toole AJ, Edelman B (2005) DISTATIS: the analysis of multiple distance matrices. Proceedings of the IEEE Computer Society: International Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA, pp 42–47

Abdi H, Valentin D, Chollet S, Chrea C (2007) Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. Food Qual Prefer 18:627–640

Aubry V, Etievant P, Sauvageot F, Issanchou S (1999) Sensory analysis of Burgundy Pinot noir wines: a comparison of orthonasal and retronasal profiling. J Sens Stud 14(1):97–117

Ballester J, Dacremont C, Le Fur Y, Etiévant P (2005) The role of olfaction in the elaboration and use of the Chardonnay wine concept. Food Qual Prefer 16(4):351–359

Ballester J, Patris B, Symoneaux R, Valentin D (2008) Conceptual vs. perceptual wine spaces: does expertise matter? Food Qual Prefer 19(3):267–276

Chollet S, Valentin D (2000) Le degré d'expertise a-t-il une influence sur la perception olfactive? Quelques éléments de réponses dans le domaine du vin [Expertise level and odour perception: what can we learn from red Burgundy wines?]. L'Année Psychologique 100:11–36

Chollet C, Valentin D (2001) Impact of training on beer flavor perception and description: are trained and untrained subjects really different? J Sens Stud 16:601–618

Chollet S, Valentin D (2006) Impact of training on beer flavour perception. Cerevisia 31(4):189–195

Clapperton JF, Piggott JR (1979) Flavour characterization by trained and untrained assessors. Journal of Institute of Brewing 85:275–277

D’Eer M (1998) La bière: ales, lagers et lambics [Beer: ales, lagers and lambics]. BièreMag, Saint-Laurent, Québec

D'Eer M (2005) Atlas mondial de la bière [World atlas of beer]. Trécarré, Outremont, Québec

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, New York

Escouffier Y (1973) Le traitement des variables vectorielles [Analysis of vector variables]. Biometrics 29:751–760

Feldon DF (2007) The implication of research on expertise for curriculum and pedagogy. Educ Psychol Rev 19:91–110

Glover B (2001) Le grand livre de la bière [The complete guide to beer]. Manise, Genève

Hughson AL, Boakes RA (2002) The knowing nose: the role of knowledge in wine expertise. Food Qual Prefer 13(7–8):463–472

Ishii R, Kemp SE, Gilbert AN, O'Mahony M (1997) Variation in sensory conceptual structure: an investigation involving the sorting of odor stimuli. J Sens Stud 12:195–214

Jackson M (1999) La bière [Beer]. Gründ, Paris

Labbé D, Rytz A, Hugi A (2004) Training is a critical step to obtain reliable product profiles in a real food industry context. Food Qual Prefer 15(4):341–348

Mastrojanni M (1999) Guide de l'amateur de bière [Guide for beer enthusiast]. Solar, Paris

Morot G, Brochet F, Dubourdieu D (2001) The color of odors. Brain Lang 79:309–320

Pangborn RM, Berg HW, Hansen B (1963) The influence of color on discrimination of sweetness in dry table wine. Am J Psychol 76:492–495

Solomon GEA (1997) Conceptual change and wine expertise. J Learn Sci 6:41–60

Valentin D, Chollet C, Béal S, Patris B (2007) Expertise and memory for beers and beer olfactory compounds. Food Qual Prefer 18 (5):776–785

Vicente KJ, Wang JAH (1998) An ecological theory of expertise effects in memory recall. Psychol Rev 105(1):33–57

Webb T (2005) Good beer guide Belgium. CAMRA Campaign for Real Ale, St. Albans