Food Quality and Preference xxx (2011) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples

R. Symoneaux [a,*], M.V. Galmarini [b,c], E. Mehinagic [a]

[a] LUNAM Université, Groupe ESA, UPSP GRAPPE, 55 rue Rabelais, BP 30748, 49007 Angers Cedex 01, France
[b] Facultad de Ciencias Agrarias, Pontificia Universidad Católica Argentina, Cap. Gral. Ramón Freire 183 (CP1429), Ciudad de Buenos Aires, Argentina
[c] Member of Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 19 May 2011
Received in revised form 30 August 2011
Accepted 30 August 2011
Available online xxxx

Keywords:
Comment analysis
Chi-square per cell
Apple
Internal preference mapping
Preference sensory key drivers

ABSTRACT

This study compares the analysis of consumer's comments resulting from a hedonic test as an alternative to the traditional internal preference mapping. During a consumer test, 87 apple consumers first evaluated six different Golden apple batches on a hedonic scale and then answered to the non-mandatory open-ended questions stating separately what they liked and disliked from each batch. In parallel, an expert panel described the sensory profiles of the studied products.

To compare the results obtained by the two studied methods the RV coefficient was calculated and was found to be 0.8656 ($p = 0.011$). Therefore, the information obtained by the comment analysis of likes and dislikes was similar to that resulting from sensory characterization done by the trained panel. With both methods, crunchiness and sweetness appeared as main sensory preference key drivers, while mealiness was not appreciated. At the same time, some characteristics such as juiciness appeared important for consumers but it was not a significant discriminant attribute for the trained panel.

A new method, the Chi-square per cell, was used to deeply analyze the contingency table of the main modalities used by consumers allowing the identification of the significant modalities which described each apple liking. Finally, the distinction between likes and dislikes made the transcription of consumers' opinions easier, without a need of interpretation on behalf of the transcoder.

1. Introduction

In the past years, food product development has become a consumer oriented task making the understanding of consumers' preference a key factor for success. Therefore, knowing how much consumers like a product is not enough, it is important to understand the reason of their choices (Chrea et al., 2010), which is a quite challenging task for both marketing and sensory scientists.

A traditional approach used by sensory science to understand consumers' preference consists in the development of internal and external preference mapping (Frewer, Risvik, & Schifferstein, 2001; Greenhoff & MacFie, 1994) combining descriptive data provided by a trained panel with hedonic tests carried out by consumers. This method has been widely and successfully used; however, the use of a trained panel can be expensive and time consuming for the industry sector. In addition, vocabulary used by the trained panel may differ from that used by consumers (ten Kleij & Musters, 2003) and also the trained panel description may be focused on attributes which are not that important for consumers (Ares, Giménez, Barreiro, & Gámbaro, 2010).

ten Kleij and Musters (2003) proposed text analysis of open-ended questions as a complementary method to the preference mapping and applied this methodology during a consumer test on mayonnaises. These authors allowed consumers to voluntarily write down remarks after their evaluations finding in this way similar results as when using internal preference mapping. However, in this case, consumers did not point out if what they mentioned was a positive (like) or negative (dislike) characteristic of the product; this subjective categorization of the comments was later assumed by the data analyst. Ares et al. (2010) did a similar work on milk desserts where commenting was compulsory for consumers. However, they were not asked separately regarding their likes and dislikes, they only had to provide up to four words to describe the product.

Asking consumers to explain why they like or dislike a product could be considered controversial. According to Lawless and Heymann (1998) consumers respond to the product as a whole, even when sometimes one or two salient characteristics may drive their decisions. In addition, these characteristics may cause other attributes to be viewed as more positive ("halo effect"). It is for this reason that these authors consider that it is difficult to get consumers to accurately explain the basis of their choice and there is a risk in asking consumers the reason for their preferences. Nonetheless,

* Corresponding author. Tel.: +33 2 41 23 55 55; fax: +33 2 41 23 55 00.
E-mail address: r.symoneaux@groupe-esa.com (R. Symoneaux).

0950-3293/$ - see front matter
doi:10.1016/j.foodqual.2011.08.013

it is important to point out that consumers are becoming more demanding and more aware of what they expect to find in certain products (Clemons, 2008). Therefore, it is our opinion that consumers' comments should not be underestimated but carefully analyzed. For this purpose, comment analysis is a technique which is becoming widely used.

Comment analysis involves counting the frequency of mention of the terms used by consumers to describe a product, obtaining then a contingency table. This data can be analyzed with statistical tools such as network segment (Rostaing, Ziegelbaum, Boutin, & Rogeaux, 1998) or, more often, correspondence analysis (CA) (Giboreau, Navarro, Faye, & Dumortier, 2001; Grunert et al., 2001; Perrin & Pagès, 2009; Sauvageot, Urdapilleta, & Peyron, 2006; Soufflet, Calonnier, & Dacremont, 2004). This technique, developed by Benzecri (1976, 1980), is a descriptive analysis of a two-way table containing measurements of correspondence. Data can finally be visualized in graphic representation (Blasius & Greenacre, 1994; ten Kleij & Musters, 2003) similar to that obtained by other multidimensional techniques (i.e. principal components analysis, discriminant analysis, and multiple factorial analysis). ten Kleij and Musters (2003) and Ares et al. (2010) conducted CA in order to analyze consumers' answers of open-ended questions after an hedonic test and found that the resulting CA map corroborated preference mapping results. Nevertheless, a deep analysis of the contingency table was not done in this study. Ares et al. (2010) carried out a global Chi-square test to study the independence between rows and columns but did not use statistical analysis to identify significant differences among products and modalities within the contingency table.

The aims of the present work were to analyze the use of comment analysis of likes and dislikes stated separately by consumers in comparison to data obtained by internal preference mapping and to deepen the text analysis by improving the statistical analysis of the contingency table. This new approach was applied on apples being the highly familiar basic product for French consumers since it is the most produced and consumed fruit in France (ANPP, 2011; FNPF, 2007).

2. Materials and methods

2.1. Samples

Six batches of Golden apples (identified with letters A through F) were selected from different regions in France, in order to introduce variability between samples. They were all purchased from wholesalers at the *Marché d'Intérêt National* (MIN) at Rungis, France, and stored at 4 °C. Before tasting, apples were washed and tempered at room temperature (21 °C ± 1 °C).

2.2. Sensory evaluation by a trained panel

Thirteen assessors (six males and seven females; 32–56 years old) of the permanent trained panel from the *Ecole Supérieure d'Agriculture* (ESA) in Angers, France, analyzed the six apple batches by quantitative descriptive analysis (QDA) method (Stone & Sidel, 1993). This panel had been trained for 4 years on apple descriptive characterization according to AFNOR recommendations (1995) being experienced on this product category (Mehinagic, Royer, Symoneaux, Bertrand, & Jourjon, 2004; Mehinagic, Royer, Symoneaux, Jourjon, & Prost, 2006).

Twelve sensory attributes were analyzed on a continuous scale: odor intensity, aroma intensity, sourness, sweetness, astringency, touch firmness, roughness, crunchiness, chewiness, juiciness, mealiness, and melting. The definitions used for these attributes are specified in Table 1. One unpeeled apple of each batch was presented to each assessor, in a sequential monadic way, according to orders based on a William Latin-square arrangement. The amount of samples was a limiting factor in this experiment so no replication was done. Nevertheless, the panel performance was tested during the training period on three different batches of Golden cultivars and the discrimination and the agreement between assessors were high. Scores for all attributes were collected with FIZZ (version 2.10; Biosystems, Courtenon, France).

All the analyses were done in individual computerized booths according to NF ISO 8589 norms. The sensory room was kept at 21 ± 1 °C, white lights were used and rinsing with mineral water between samples was mandatory.

2.3. Consumers test

A total of 87 apple consumers recruited in the city of Angers, France, participated in the test. Gender of participants was balanced, being 54% female and 46% male, and ages ranged from 18 to 60 years old. Whole unpeeled apples were presented in the same way as for the sensory panel (sequential monadic way) and consumers were asked to rate their liking on a seven-point hedonic scale. The hedonic scale ranged from 0 to 7 being 0 an extremely disliked apple and 7 an extremely liked one. It is to be noted that, the hedonic scale method was used under the special consideration of removing the neutral category (ASTM, 1968) in order to allow a clearer correlation between ratings and consumer's comments ("likes" and "dislikes"). It is known that this kind of variations in the scale can cause marked changes in the distribution of responses and statistical parameters such as means and variances; however, relative measures tend to remain constant (ASTM, 1968).

After expressing their level of overall liking, consumers were given the option to freely state separately, what they liked ("likes", L) and what they disliked ("dislikes", D) about each sample. Answering both open-ended questions was not mandatory. In this way, they could express only likes, only dislikes, both or none for each product.

2.4. Data analysis

2.4.1. Analysis of variance (ANOVA)

Two-ways analysis of variance (ANOVA) were conducted to assess significant differences between apple batches in terms of sensory attributes and overall liking scores considering product and tasters as fixed sources of variation. A significance level of 5% was considered. The mean intensities were then compared by Student Newman–Keuls (SNK) multiple comparison test. Statbox software (Version 6.6, Grimmersoft, Issy les Moulineaux, France) was used to perform this analysis.

2.4.2. Principal component analysis of trained panel data

A principal component analysis (PCA) was performed with R language (R Development Core Team., 2011) and FactoMineR (Husson, Bocquet, & Pagès, 2004) only on the significant attributes identified by ANOVA and averaged per product. A correlation matrix was used and the minimum Eigen value was set at 1.

2.4.3. Internal preference mapping

A principal component analysis (PCA) of the correlation matrix of consumers (variables) by products (objects) was carried out with FactoMineR (Husson et al., 2004). Sensory attributes were added as supplementary variables.

2.4.4. Text analysis

Text analysis requires a shaping or structuring of the data since consumer's responses are not exploitable in their initial state. Each consumer wrote his/her comments without guidance, in a personal

Table 1
Definition of the attributes used by the trained panel to describe the apples.

Attribute	Definition
Flavor	
Odor intensity	Strength of the external odor in the apple sample (nasal aroma)
Aroma intensity	Aroma released during chewing (retronasal aroma)
Sourness	Taste related to acids (e.g. malic acid in apples)
Sweetness	Taste related to simple sugars (e.g. sucrose, fructose)
Astringency	The shrinking or puckering of the tongue surface caused by substances such as tannins and potassium aluminum sulfate
External touch sensations	
Touch firmness	Resistance of fruit to pressure applied with thumb and index fingers
Roughness	Degree of apple peel's roughness as measured by touch
Texture	
Crunchiness	The combination of the force required for the first bite and the noise resulting from this bite
Chewiness	Time and number of chewing movements needed to grind the sample prior to swallowing
Juiciness	Amount of liquid released during mastication
Mealiness	Degree to which the flesh breaks down to a fine lumpy mass and to very fine dry particles
Melting	Force required to crush a piece of unpeeled apple between the tongue and palate

style, even with typing faults, orthographic and grammatical mistakes; and these needed to be transformed into precise modalities.

Rostaing et al. (1998) precise the procedure of postcoding as: removing mistakes, elimination of connectors and auxiliary terms, location of phrases and terms which make them up, lemmatization, regrouping synonyms, managing ambiguous words (polysemy and homographs), marking terms specific to sensory analysis.

Table 2
Example of the transformation of free comments into structured modalities. "Likes" for the first ten consumers for batch C are shown.

Raw likes	
C1	Lots of taste, crisp, quite sweet
C2	Distinct apple odor not much taste, somewhat sweet
C3	Juicy, sugary, slightly acidic, crisp, cellar odor
C4	Odor
C5	Quite juicy
C6	
C7	Good apple
C8	
C9	Juicy, crispy
C10	Juicy, crispy, quite acid, good taste, good texture
Likes after simplification	
C1	L_Taste; L_Crisp; L_Sweet
C2	L_Odor; L_Sweet
C3	L_Juicy; L_Sweet; L_Sour; L_Crisp; L_Cellar Odor
C4	L_Odor
C5	L_Juicy
C6	
C7	
C8	
C9	L_Juicy; L_Crisp
C10	L_Juicy; L_Crisp; L_Sour; L_Taste; L_Texture
Likes count (number of citations)	
L_Sour	2
L_Crisp	4
L_Taste	2
L_Juicy	4
L_Odor	2
L_Cellar Odor	1
L_Sweet	3
L_Texture	1

In the present manuscript, the dataset was presented in a MSExcel file having for each consumer: (a) the product, (b) all the initial information provided by the consumer for likes, and (c) all the initial information provided by the consumer for dislikes; all these in separate rows. The transcoder had to: (1) verify typing and/or spelling mistakes, (2) add a new column where all connectors, auxiliary terms and adverbs were deleted from each comment in a second column, and (3) regroup terms from this new column (lemmatization). Then taking into account the initial complete full comment, it was possible to distinguish modalities with the same base and nuance.

Particularly in this database, synonyms were quite evident; therefore the transcoder could make the regrouping. In case of ambiguous comments, no action was taken and all terms were kept without any regrouping. Therefore, the two most delicates steps (regrouping synonymous and managing ambiguous words) were done in a way that no over interpretation or over grouping of modalities took place. During this simplification process, all liking comments were transformed in simpler modalities beginning by "L_" and all disliking ones by "D_".

As an example, in Table 2, all the likes for apple C for the first 10 consumers are presented, enabling the visualization of how free comments are transformed into more structured modalities. The first stage for data simplification was the removal (as for Judge 7, Table 2) of hedonic terms (good apple, nice) since the purpose of the work was to find words which could be related to sensory attributes. Given that questions already considered separately likes and dislikes; hedonic terms did not add relevant information for this aim.

Then, likes and dislikes were re-transcribed into simpler modalities for each consumer and each apple. For example, for consumer 10, who noted down as likes of apple C: "juicy, crisp, slightly acid, pleasant taste; good texture", the following modalities were transcribed: L_Juicy; L_Crisp; L_Acid; L_Taste; and L_Texture (Table 2).

Once the re-transcription of the 87 consumers for the six different batches and for likes and dislikes was done, modalities per product were counted. Table 4 presents the contingency table for the total of like and dislike comments and those modalities mentioned at least by 5% of the consumers for one product or more.

2.4.5. Global Chi-square and Chi-square per cell

Global Chi-square is used for testing the independence between rows and columns of the contingency table. When the initial Chi-square is significant it is possible to analyze within each cell identifying the source of variation of the global Chi-square (Snedecor & Cochran, 1957). In the present work, this analysis was done using the specific software Statbox (Version 6.6, Grimmersoft, Issy les Moulineaux, France). An excel file with the columns judge, product, modalities for likes and modalities for dislikes was used. The contingency table was produced using the function "Cross tabulation" with batches as columns and modalities as rows. Then by means of the command «Chi-square per cell» a table was obtained, showing for each cell of the contingency table if the observed values of each cell are significantly higher, lower or equal to the theoretical values. This test is performed on a 2×2 table where one cell is a $[i,j]$ cell of the original contingency table and the others contain values for the row i minus $[i,j]$, for the column j minus $[i,j]$ and for the rest of the table. A Chi-square test is then performed for each 2×2 table.

2.4.6. Correspondence analysis

In order to visualize the relationship between products and likes and dislikes cited by consumers, correspondence analysis was performed on the contingency table (product vs. L and D items) on the main modalities that at least 5% consumers used

Table 3
Mean scores of sensory attributes and consumer preference for each sample batch (A–F).

Attribute	Apple batches						*p* Value
	A	B	C	D	E	F	
Odor intensity	3.01 bc	2.30 ab	2.31 ab	4.09 c	1.76 a	3.38 bc	0.0020
Aroma intensity	4.05 ab	4.88 b	3.55 a	4.70 b	4.79 b	3.14 a	0.0048
Sourness	2.60 ab	3.56 c	2.79 abc	1.93 a	2.92 bc	3.16 bc	0.0185
Sweetness	6.01 b	6.35 b	6.03 b	6.27 b	6.50 b	4.78 a	0.0064
Astringency	2.29 a	2.89 a	2.81 a	2.44 a	2.97 a	2.87 a	0.5540
Touch firmness	5.60 a	7.49 b	6.31 a	5.71 a	7.45 b	5.66 a	<0.0001
Roughness	1.10 ab	2.12 c	1.03 ab	1.73 bc	0.80 a	0.65 a	0.0032
Crunchiness	5.10 a	7.38 b	5.97 a	5.35 a	7.15 b	5.64 a	<0.0001
Chewiness	5.61 a	6.72 bc	5.78 a	5.48 a	6.92 c	6.03 ab	0.0027
Juiciness	4.38 a	4.16 a	4.61 a	3.95 a	5.13 a	4.92 a	0.2564
Mealiness	3.75b	2.06 a	2.56 ab	3.87 c	2.67 abc	2.18 a	0.0164
Melting	6.30 d	2.55 a	4.36 bc	5.28 cd	2.48 a	4.05 b	<0.0001
Overall liking score	3.24 b	4.25 c	3.67 b	2.55 a	4.30 c	3.72 b	<0.0001

Different lower case letters represent significant differences ($p < 0.05$) among samples according to Student Neuman–Keuls (SNK).

Table 4
Contingency table showing main modalities cited by consumers and the total number of like and dislike modalities. Number of citations per batch and results of the chi square per cell are presented.

Main modalities[a]	Apple batches					
	A	B	C	D	E	F
Like comments (L)						
Crunchy	7 (−)***	32 (+)***	18	5 (−)***	28 (+)**	20
Sweet	17	14	14	14	26 (+)**	7 (−)**
Juicy	17	13	25 (+)*	13	15	23
Taste	4 (−)*	18 (+)***	10	4 (−)*	12	10
Firm	6	13	7	2 (−)**	16 (+)**	8
Sour	0 (−)**	9	8	3	9	7
Texture	3	3	9 (+)***	3	3	1
Coulor	4	2	4	7	8	3
Aspect	1	1	5	7	2	6
Soft	3	2	2	7 (+)*	3	3
Sweet/Sour_Ratio	2	5	2	0	4	2
Soft_Flesh	3	1	1	5 (+)**	0	1
Dislike comments (D)						
Mealy	10	1 (−)***	11	19(+)***	4 (−)*	8
Tasteless	27 (+)**	9 (−)**	16	26 (+)**	14	16
Not_Crunchy	13 (+)***	1 (−)**	5	11 (+)*	2 (−)*	6
Not_Sweet	11	3 (−)**	13	10	5	13
Taste	4	2	3	12 (+)***	4	5
Sour	4	12 (+)***	7	2	2	5
Soft_Flesh	9 (+)**	1	2	8 (+)*	1	5
Thick_Skin	2	3	1	4	7	5
Hard_Skin	0 (−)*	5	6	3	6	1
Hard	0	6 (+)**	1	2	5	0
No_Odor	4	2	6	3	6	6
Firm	0	5 (+)**	1	1	5 (+)*	0
Aspect	0	1	1	0	5 (+)**	4
Modalities						
Total like comments (L)	96 (−)***	140(+)***	125	86 (−)***	147(+)***	113
Total dislike comments (D)	133 (+)***	96 (−)***	115	153(+)***	109(−)***	115

(+) or (−) indicate that the observed value is higher or lower than the expected theoretical value.
[a] Modalities mentioned at least by 5% of the consumers for at least one product.
* $p \leqslant 0.05$; effect of the Chi-square per cell.
** $p \leqslant 0.01$; effect of the Chi-square per cell.
*** $p \leqslant 0.001$; effect of the Chi-square per cell.

for at least one product. The correspondence analysis (CA) maps the data generated from a contingency table representing rows and columns in the same geometric space. This analysis was done using SPAD software (SpadVersion: MN:6.5.0, Paris, France).

2.4.7. RV coefficient

Finally, the regression vector (RV) coefficient (Perrin & Pagès, 2009) calculated using the function CoeffRV from FactoMineR (Husson et al., 2004) was also computed between the first five axes of the CA and of the PCA to identify the link between both methodologies and to analyze similarity between both characterizations.

3. Results

3.1. Trained panel data analysis

ANOVA results for the sensory panel description of the six batches are presented in Table 3. This shows that apple batches were different ($p < 0.05$) for the attributes describing most flavor and texture attributes, except for juiciness and astringency. The principal component analysis was carried out on the different ($p < 0.05$) mean values obtained for each batch and it is presented in Fig. 1.

The first two components explained 59% and 27% (components 1 and 2, respectively) of the total variation. The first component (PC1) was principally explained by firmness and crunchiness. Apples from batches B and E which were characterized mainly by these two attributes (see also Table 3) had a high contribution on this component. Moreover, the attributes melting and odor intensity, opposite to the aforementioned, had also an important weight on this component.

The second component (PC2) distinguished batches based on roughness, aroma intensity, mealiness, sweetness, and sourness. In this way, batch F appeared separated from the rest, because of its low sweetness intensity while batches B and D had a higher roughness and aroma intensity (this last attribute was also high for batch E; Table 3).

The third component (not shown) was responsible for 11% of the variation and explained differences mainly in terms of roughness, allowing the separation of samples B and E, since sample B was significantly rougher than sample E. In addition, it could be seen that in terms of sourness sample B is somewhat sourer than E.

Fig. 1 also reveals that the relationship between odor (nasal aroma) and aroma (retronasal aroma), as described by the trained panel, was independent. The same can be observed for the measurement of the basic tastes sweetness and sourness. However, sweetness and aroma intensity were highly related. This is supported by the known fact that sweetness can somehow enhance the aromatic perception of fruity compounds (Noble, 1999).

3.2. Preference mapping

Table 3 shows the mean values for the overall liking of the 87 consumers. Ratings were different ($p < 0001$) among samples showing that batches B and E had the highest preference. Batches A, C, and F presented intermediate liking scores while the batch D was the least appreciated by consumers.

Fig. 2 shows the internal preference map of consumers' hedonic scores. The first two principal components (PC) represented 51% of the total variation. The dispersion of consumers on the right of the graph indicated some common liking pattern between consumers

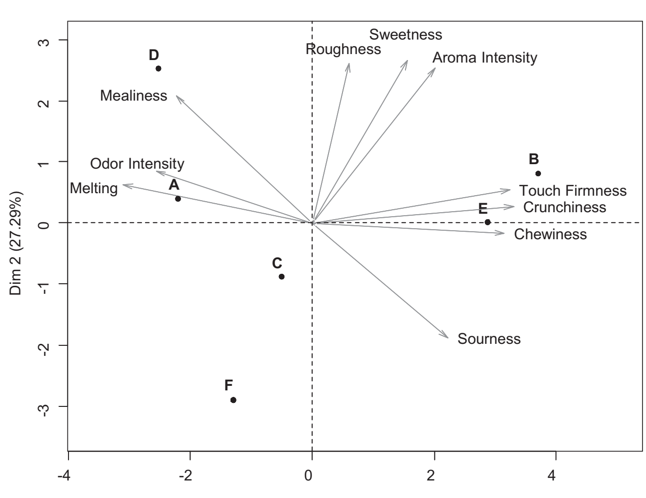

Fig. 1. Principal component analysis of the significant sensory attributes mean values ($p < 0.05$) of the trained panel's data.

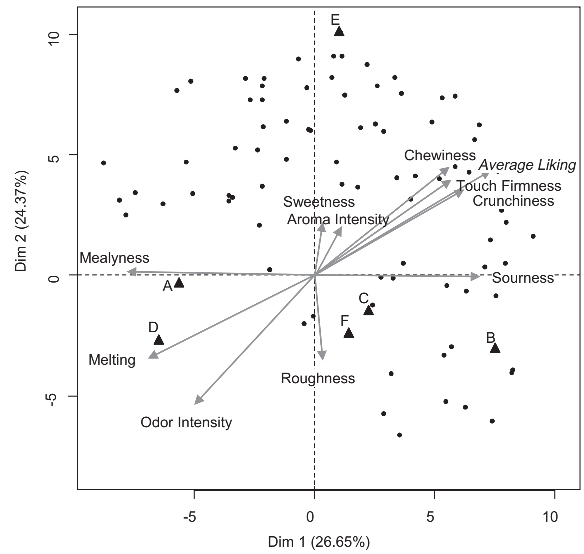

Fig. 2. Internal preference mapping of consumers' hedonic scores (black points) with projection as supplementary variables of the trained panel data (grey vectors).

on the first PC but some differences at the same time on the second one. The location of products A and D on the graph confirmed the low preference of these two batches. Batches C and F were intermediately appreciated, being close to the center of coordinates. The high positive contribution on PC2 of batch E, opposed to the location of batch B, showed that even if these two batches had a high consumer's acceptability, they were not chosen by the same consumers, those who preferred B did not choose E.

The projection of sensory attributes on internal preference mapping allowed the identification of some sensory key-drivers of preference. The attributes mealiness and melting had a negative contribution on PC1 and were opposite to consumer preference. Therefore, batches A and D, which had high scores for mealiness and melting, had a low preference. These results are in accordance with a previous work by Jaeger, Andani, Wakeling, and MacFie (1998) which showed that consumers consider mealiness in apples as a negative quality attribute, associated with granular and floury texture.

The attributes which were highly related in a positive way to preference were sourness, firmness, and crunchiness. The high scores for firmness and crunchiness given to batches B and E were probably the ones which accounted for their higher preference. However, their difference in roughness (Table 2) and the slight difference in sourness resulted in variation among consumers' liking for each batch. These results are in accordance with those of Daillant-Spinnler, MacFie, Beyts, and Hedderley (1996) where an apple consumers' segmentation was found based on whether an apple was sweet and hard or juicy and sour.

Finally, it is to be noted that odor intensity seemed to have a negative correlation with preference.

3.3. Comment analysis on likes and dislikes

Consumers were asked two open-ended questions to know what they liked and/or disliked about each batch after they had given their overall liking score. Table 4 shows the main modalities for liking (L) and disliking (D) that consumers used.

The most recurrent likes mentioned by consumers were crunchy, juicy, and sweet. These results are in accordance with a large survey on French apples consumers carried out by Hutin (2008) which showed that most French consumers look first for crunchy

Please cite this article in press as: Symoneaux, R., et al. Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. *Food Quality and Preference* (2011), doi:10.1016/j.foodqual.2011.08.013

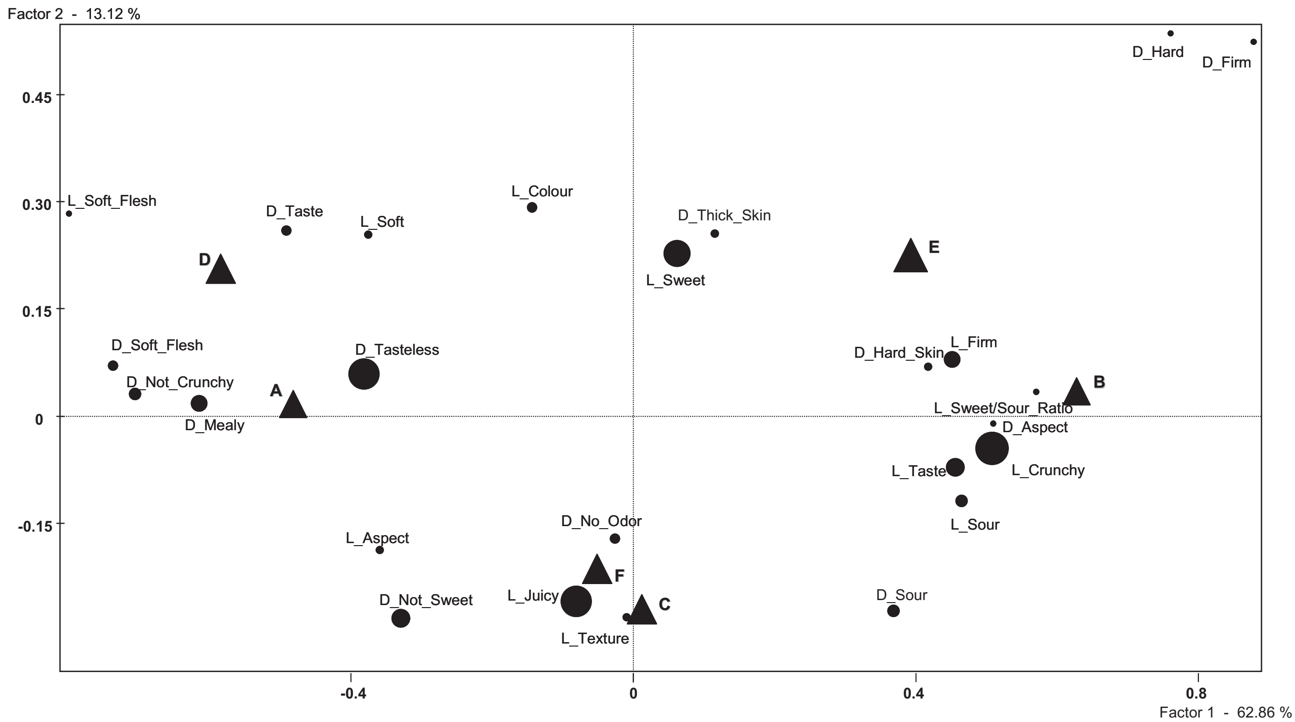

Fig. 3. Correspondence analysis on contingency table crossing apple batches and main modalities cited by consumers. Triangles represent products and circles represent modalities. The size of both triangles and circles is a function of the contribution of each product or modality in the correspondence analysis.

and sweet apples. On the other hand, the principally mentioned dislikes were tasteless, mealy, not crunchy and not sweet.

The number of like and dislike comments obtained were related to consumers' overall liking ratings by a Chi-square test (χ^2 = 262.96, $p < 0.001$). It was found that the higher the ratings given for a product, the more comments consumers gave for likes and the less they gave for dislikes. Also, lower ratings were related to more dislike than like comments. Batches B and D, which had the highest and lowest overall likings respectively (Table 3) had 140 and 86 like comments and 96 and 153 dislike comments, respectively. On the other hand, products with an intermediate preference such as C and F received the same amount of like and dislike comments (see Table 4). In this way consumers' liking score were consistent with their liking/disliking comments. The more appreciated a product was, the more liking comments it received and vice versa. Neutral products in terms of liking received the same amount of liking and disliking comments.

Global Chi-square carried out on the contingency table (Table 4) was highly significant (χ^2 = 931.17, $p < 0.001$) revealing that apple consumers used different modalities for describing their likes and dislikes depending of each batch.

The Chi-square per cell allowed identifying which type of modality (L or D) was more or less used for one or several products. In Table 4 it can be observed that batch B obtained more crunchy (L), taste (L), and sour (D) citations than any other batch and less tasteless (D) and mealy (D). Batch E had a high frequency of mention for crunchy (L) just as batch B, but was also characterized with a higher proportion of sweet (L) and firm (L) modalities.

Batch D, as well as A, had the lowest frequency of mention for firm (L) and crunchy (L) modalities. As dislikes for this batch, the most mentioned by consumers were mealy (D), tasteless (D), and taste (D). The characterization of sample A was similar to D in terms of dislikes too. It also had high frequency of mention for the modalities tasteless (D), not crunchy (D), and soft flesh (D).

Batch C was positively characterized by the modalities juicy (L) and texture (D) while it did not present any dislikes which differentiated it from the other batches. Finally, batch F was set apart from the rest by being described as the least sweet (sweet (L)).

In order to visualize the characterization of the product based only on the comments cited by consumers, a correspondence analysis (CA) was carried out; results are presented in Fig. 3.

The two first dimensions of the CA represented 76% of the total variation. According to comment analysis, products were clearly separated along the first axe into the same three groups as those formed by analysis of overall liking ratings (Table 3). Batches D and A were on the left, F and C in the middle and B and E on the right. Batches D and A were disliked because they were mealy (D), less crunchy (D) and with a soft flesh (D). Opposite to these, batches B and E were liked for being firm (L), crunchy (L) and for its sweet/acid ratio (L).

The analysis of the third component in the CA (not shown) and the contingency table allowed identifying three clusters; batches D and A, batches F and C and finally batches E and B. In this last batch, some consumers found apples B to be too sour (dislike characteristic) and liked batch E because of its higher sweetness. This is in accord with the information obtained by the sensory profile for each batch. On the other hand, the higher roughness mentioned for B by the trained panel is not mentioned by consumers as a liking or disliking characteristic. Nevertheless, it appeared that some consumers mentioned that they disliked the aspect for apples E. These could explain some nuances in consumers liking.

4. Discussion

Batches D and A presented the highest mean value for the attribute mealiness by the trained panel (see Table 3) and then consumers stated that they disliked this apple because it was mealy (see Table 4) and not crunchy. It is to be noted that in terms of likes consumers always mentioned presence of something they considered positive; the absence of a defect was never considered as a positive characteristic. On the other hand, when talking about dislikes, they mentioned either the presence of a characteristic they disliked or the absence of something they expected (e.g. not_sweet, Table 4).

However, it is to be noted that some nuances could be observed between the words used by consumers and the trained panel. For example, for batches A and D the most used dislike modality was tasteless. Meanwhile, the trained panel described these batches as the least sour while having similar sweetness. Therefore, a hypothesis could be that, when consumers say that they dislike a sample because it is tasteless, they actually mean that it is not sour enough. In addition, even when the trained panel did not find any difference concerning juiciness ($p = 0.2564$), consumers distinguished batch C by the juicy (L) modality. In this way, consumers' and panel's perception of juiciness is probably not the same.

Moreover, consumers emphasized on attributes which were not treated by the trained panel. For example, in the dislike modalities, consumers referred to the apple skin (thick skin, hard skin); attribute which did not seem pertinent of evaluation to the panel leader.

The use of open-ended questions to analyze consumer's sensory key drivers was suggested by ten Kleij and Musters (2003) and improved by Ares et al. (2010). In both cases consumers were asked to write comments after giving their liking note. ten Kleij and Musters (2003) asked to write down remarks to explain why they gave particular liking scores while Ares et al. (2010) demanded consumers to provide up to four words to describe each sample. However, neither of them asked consumers to state whether descriptors were positive or negative, leaving this discrimination task to the transcoder. In the present work, consumers were asked to comment after giving their liking score distinctly stating what they liked and what they disliked about each product. In this way, a clear separation between liking and disliking key drivers was obtained; without further need of interpretation. As consumers made this discrimination, the transcription was easier, leading to a better understanding of consumers' comments. Often, when consumers answer to only one open-ended question, it can be hard to know if what they mention is for them a positive or negative aspect. For instance, if consumers had said only "aspect" or "firm" the obtained information would have been very limited, since it is not possible to know if they liked this or not. When asked separately, not only does this interpretation become easier; but it can be found that a same characteristic can be liked by some consumers and disliked by others. That was the case for aspect, soft flesh, firm, and sour (Table 4).

Moreover, it was found that the amount of liking and disliking comments given was related to the overall liking score of each batch. For example, for low overall liking scores (between 0 and 2), consumers gave more disliking than liking comments (around 20 L for 80 D). For high ratings (5–7), the opposite was observed (80 L for 20 D) while when grading with 3–4, consumers gave the same amount of liking and disliking comments. In this way, it was found that consumers are more loquacious about the products they like when asked to distinguish between liking and disliking comments as opposed to that reported by ten Kleij and Musters (2003) who observed that consumers were less productive for a product they liked when asked only a global question. Therefore, we think that this distinction in the way consumers are inquired allows obtaining more information and a better description of the studied products.

The use of Chi-square per cell led to an improvement in the statistical evaluation of the comment analysis data. This allowed a deeper analysis of the contingency table of comments because the identification of the modalities more or less cited by consumers for one or several products could be compared to the other ones. Usually people just make a large analysis of the contingency table without focusing on this point and without having an inferential test to help users describe results. As for the use of SNK test or a least square difference (LSD) test after Fisher test in a quantitative approach, this methodology allowed to focus on the differences between products and on the observation of the significant deviation in citation between products and modalities. This methodology allows identifying the significant modalities inside the matrix and is an interesting tool to analyse the correspondence analysis results. Indeed, it could be used to know if some particular positioning of modalities on the map could be consider as significant or just as an artifact of the multidimensional calculation. For instance, from Fig. 3 it could be interpreted that consumers used the modality no_odor to describe why they disliked batches F and C. However, Table 4 shows that the difference with other products in terms of this modality is not significant. In this way, this analysis is helpful to not misinterpret results from CA mapping.

Limitations on the use of open-ended questions were mentioned by Lebart and Salem (1994) who explained that its three main drawbacks were: transcoder's mediation, deconstruction of the form and impoverishment of the meaning because uncommon answers are eliminated a priori. Nevertheless, the consistence between comment analysis and QDA profile demonstrated the pertinence of this approach. Even if the impact of the transcoder was reduced by asking separately about likes and dislikes, it can not be ignored. In marketing studies, this transcription from complete text into simpler items is often done by several people. In this study, one transcoder specialist in consumer science and apple perception was used to reduce the time of transcription's process. The comparison between different transcoders with different level of implication and knowledge of sensory science could be studied in further works.

Analyzing the number of comments given by consumers a new possible improvement arises. It was seen that around 50% of the consumers gave three comments or more (L and D) per product. But due to the non mandatory state of the questions, some consumers (6% approximately) did not answer either for like or dislike comments (Table 5). The methodology proposed by Ares et al. (2010) to ask in a mandatory way could be interesting to use. However, some consumers could have no liking comments if they totally dislike a product or have no disliking comments if they completely like it. In deed, we observed that depending on the product appreciation, from 21% to 39% of the consumers gave no like comments and from 9% to 29% gave no dislike comment, depending on the batch. So we consider that a viable possibility could be to demand at least one like or dislike comment per product, as a compromise, in order to increase the number of citations having a better description of the product.

Familiarity with the product and the complexity of the product could impact the pertinence of this methodology. Merlo and

Table 5
Number of consumers per batch writing a given number of comments.

	Apple batches					
	A	B	C	D	E	F
Number of total comments (L and D) per batch						
0	6	6	6	5	5	6
1	12	10	9	12	13	14
2	29	25	26	21	23	24
3 or more	40	46	46	49	46	43
Number of liking comments (L) per batch						
0	33	18	21	34	20	19
1	25	26	26	25	22	26
2	19	24	26	23	22	16
3	10	19	14	5	23	16
Number of disliking comments (D) per batch						
0	14	25	21	8	20	16
1	31	39	35	32	39	39
2	28	17	18	28	18	22
3	14	6	13	19	10	10

Please cite this article in press as: Symoneaux, R., et al. Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. *Food Quality and Preference* (2011), doi:10.1016/j.foodqual.2011.08.013

Mansur (2004) showed in a study concerning descriptive discourse that participants expressed more attributes when the topic is familiar than when it is unfamiliar. Since in France, apple is the most consumed fruit (ANPP, 2011), it was probably easier for consumers to describe what they liked and what they disliked in this familiar basic product than for a less known or more complex product such as wine or processed foods. The lack of familiarity and the multidimensionality of the perception could complex consumers' answers and the comments analysis.

For these reasons and because, as aforementioned, in some cases words used by consumers are not precise enough and need a sensory profile to better understand consumers, this methodology might not replace the traditional QDA and preference mapping. However, we consider that it can become an interesting alternative if the sensory profile by a trained panel can not be done (i.e. for economic reasons). In addition, it can be a complement of the traditional approach in order to access to words used by consumers, giving interesting outputs for marketers.

In preference mapping technique, the use of hierarchical cluster analysis (HCA) to segment consumers on their appreciation can be performed on their liking scores observing consumers segmentation. The use of the trained panel results allows further identification of sensory key drivers for each group. In the case of our proposed methodology, clustering can also be done and a new contingency table per cluster can be obtained finally visualizing the key drivers of each cluster on the CA map. Nevertheless to be efficient, the clusters need to have enough consumers to obtain enough comments per cluster and a significant contingency table.

5. Conclusion

In the case of a familiar product, comment analysis of consumers' likes and dislikes resulted as an interesting alternative to preference mapping method. A high correlation was obtained between the two studied methodologies. However, some nuances as well as complementary information were found.

The use of Chi-square and Chi-square per cell on the contingency table with the number of citation per product and modalities followed by correspondence analysis was an interesting tool to analyze this type of data. The deeper analysis of the contingency table allows more accuracy on data interpretation complementing the representation of comments by CA as the classical tool used in preference mapping.

Since some nuances are observed between consumer's comments analysis and trained panel results, this methodology must be used carefully. The use of a trained panel is still important to validate differences between consumers' subjective perception and objective characterization by a trained panel. Nevertheless, in absence of a trained panel, this methodology could be useful to access consumers' sensory key drivers.

Acknowledgments

The authors are thankful to Mathilde Danard for her technical help during experiments realization, to Myriam Reveillère for her statistics help and to Maria Clara Zamora for her precious advises and support during the redaction of this paper.

References

AFNOR. (1995). Analyse Sensorielle-Recherche et Sélection de Descripteurs Pour l'élaboration d'Un Profil Sensoriel Par Approche Multidimensionnelle *Contrôle de la qualité des produits alimentaires-Analyse sensorielle* (pp. 276–310). AFNOR.

ANPP. (2011). *Pomme et... chiffres*. Available from: http://www.lapomme.org/.

Ares, G., Giménez, A., Barreiro, C., & Gámbaro, A. (2010). Use of an open-ended question to identify drivers of liking of milk desserts. Comparison with preference mapping techniques. *Food Quality and Preference, 21*(3), 286–294.

ASTM (1968). *ASTM, manual on sensory testing methods special technical publication* (Vol. 434). Philadelphia: American Society for Testing and Materials.

Benzecri, J. P. (1976). L'analyse des données, l'analyse des correspondances 2. *Editions Dunod.*

Benzecri, J. P. (1980). L'analyse des données, l'analyse des correspondances 1. *Editions Dunod.*

Blasius, J., & Greenacre, M. (1994). *Computation of correspondence analysis. Correspondence analysis in the social sciences. Recent developments and applications*. London Academic Press, pp. 53–78.

Chrea, C., Melo, L., Evans, G., Forde, C., Delahunty, C., & Cox, D. N. (2010). An investigation using three approaches to understand the influence of extrinsic product cues on consumer behavior: An example of Australian wines. *Journal of Sensory Studies, 26*(1), 13–24.

Clemons, E. K. (2008). How information changes consumer behavior and how consumer behavior determines corporate strategy. *Journal of Management Information Systems, 25*(2), 13–40.

Daillant-Spinnler, B., MacFie, H. J. H., Beyts, P. K., & Hedderley, D. (1996). Relationships between perceived sensory properties and major preference directions of 12 varieties of apples from the Southern Hemisphere. *Food Quality and Preference, 7*(2), 113–126.

FNPF. (2007). *Surfaces et volumes pour quelques produits*. Available from: http://www.fnpfruits.com/sites/fnpfweb/chiffresK/productions/default.aspx#Pomme.

Frewer, L., Risvik, E., & Schifferstein, H. (2001). *Food, people, and society: A European perspective of consumers' food choices*. Springer.

Giboreau, A., Navarro, S., Faye, P., & Dumortier, J. (2001). Sensory evaluation of automotive fabrics: The contribution of categorization tasks and non verbal information to set-up a descriptive method of tactile properties. *Food Quality and Preference, 12*, 312–322.

Greenhoff, K., & MacFie, H. J. H. (1994). Preference mapping in practice. In H. J. H. MacFie & D. M. H. Thomson (Eds.), *Measurement of food preferences*. London, UK: Chapman & Hall.

Grunert, K. G., Lahteenmaki, L., Nielsen, N. A., Poulsen, J. B., Ueland, O., & Astrom, A. (2001). Consumer perceptions of food products involving genetic modification – Results from a qualitative study in four Nordic countries. *Food Quality and Preference, 12*, 527–542.

Husson, F., Bocquet, V., & Pagès, J. (2004). Use of confidence ellipses in a PCA applied to sensory analysis application to the comparison of monovarietal ciders. *Journal of Sensory Studies, 19*(6), 510–518.

Hutin, C. (2008). La pomme auprès des français. *Infos CTIFL, 195*, 11–15.

Jaeger, S. R., Andani, Z., Wakeling, I. N., & MacFie, H. J. H. (1998). Consumer preferences for fresh and aged apples: A cross-cultural comparison. *Food Quality and Preference, 9*(5), 355–366.

Lawless, H. T., & Heymann, H. (1998). *Sensory evaluation of food: Principles and practices*. New York: Chapmann and Hall.

Lebart, L., & Salem, A. (1994). *Statistiques textuelles*. Paris: Dunod.

Mehinagic, E., Royer, G., Symoneaux, R., Bertrand, D., & Jourjon, F. (2004). Prediction of the sensory quality of apples by physical measurements. *Postharvest Biology and Technology, 34*(3), 257–269.

Mehinagic, E., Royer, G., Symoneaux, R., Jourjon, F., & Prost, C. (2006). Characterization of odor-active volatiles in apples: Influence of cultivars and maturity Stage. *Journal of Agricultural and Food Chemistry, 54*(7), 2678–2687.

Merlo, S., & Mansur, L. L. (2004). Descriptive discourse: Topic familiarity and disfluencies. *Journal of Communication Disorders, 37*(6), 489–503.

Noble, A. C. (1999). Taste-aroma interactions. *Trends in Food Science and Technology, 7*(12), 439–444.

Perrin, L., & Pagès, J. (2009). Construction of a product space from the ultra-flash profiling method: Application to 10 red wines from the Loire valley. *Journal of Sensory Studies, 24*(3), 372–395.

R Development Core Team. (2011). *R: A Language and Environment for Statistical Computing*. Retrieved from: http://www.R-project.org.

Rostaing, H., Ziegelbaum, H., Boutin, E., & Rogeaux, M. (1998). Analyse de commentaires libres par la techniques des réseaux de segments. In *Fourth International Conference on the Statistical Analysis of Textual Data, JADT'98*.

Sauvageot, F., Urdapilleta, I., & Peyron, D. (2006). Within and between variations of texts elicited from nine wine experts. *Food Quality and Preference, 17*(6), 429–444.

Snedecor, G. W., & Cochran, W. G. (1957). *Statistical methods* (6th ed.). Iowa, USA: The Iowa State University Press.

Soufflet, I., Calonnier, M., & Dacremont, C. (2004). A comparison between industrial experts' and novices' haptic perceptual organization: A tool to identify descriptors of the handle of fabrics. *Food Quality and Preference, 15*(7–8), 689–699.

Stone, H., & Sidel, J. L. (1993). *Sensory evaluation practices* (2nd ed.). San Diego: Academic Press, Inc..

ten Kleij, F., & Musters, P. A. D. (2003). Text analysis of open-ended survey responses: A complementary method to preference mapping. *Food Quality and Preference, 14*(1), 43–52.