Exploring consumer product profiling techniques and their linkage to a Quantitative Descriptive Analysis

Karima A. Moussaoui[1] and Paula Varela[2,3]*

[1]*Sensory-Consumer Science, Kraft Foods Germany*
Bayerwaldstraße 8, 81737 München

[2]*Sensory-Consumer Science, Kraft Foods UK*
Ruscote Avenue, OX163WN, Banbury, United Kingdom

[3] *Current affiliation: Institute of Agrochemistry and Food Technology - Avda. Agustín Escardino 7, 46980-Paterna- Valencia, Spain.*

*Corresponding author: Paula Varela

Email: pvarela@iata.csic.es

Tel: (+34) 963 90 0022 Ext. 2227

Fax: (+34) 963 63 6301

Abstract

Consumer's voice is crucial for new product development. One way to capture it is to ask consumers to describe products and to quantify their perception of this description. In this context four profiling methods; Sorting, Projective Mapping, Flash Profile and Repertory Grid Method (RGM) were explored among target consumers of hot beverages in two European countries (UK and France) with the assumption that meaningful sensory descriptors can be generated and quantified, and that product maps can ultimately be drawn. A Quantitative Descriptive Analysis was also performed with a trained panel and its outcomes were used as a basis for comparison. Results showed that consumers were able to describe and quantify product differences, that their perception was similar on a cross-country level,that trained panel maps translated well consumers' description, and that Flash Profiling and RGM were more suitable for such a task as they generate a rich vocabulary and more accurate maps. However, when describing complex attributes as mouthfeel or afterfeel, the consumers' description was not enough detailed or not consensual.

Keywords: New Product Development, Sorting, Projective Mapping, Flash Profile, Repertory Grid, Quantitative Descriptive Analysis, Sensory, Consumers.

1. Introduction

Sensory evaluation can be seen as a link between Research and Development, with a focus made on technical aspects of food, and Consumer and Marketing Research, with a focus on consumers' behaviour and psychology (Dijksterhuis, 1997). They measure the reaction to stimuli resulting from the use or consumption of a product through analytical and/or affective tests. Traditionally, analytical tests (discriminative and descriptive) are performed with trained panels whereas affective tests are run with consumers (Stone and Sidel, 1993).

QDA(R) method is based on the principle of a panellist's ability to verbalize perceptions of a product in a reliable manner; panellists are screened and trained in attribute recognition and scaling, they use a common and agreed sensory language, and products are scored on repeated trials to obtain a complete, quantitative description (ASTM, 1992). Describing the sensory characteristics of products has been an integral part of the food and beverage industry since long ago. Information obtained from the description of the sensory characteristics of food and beverages enable companies to make more informed business decisions (Stone and Siedel, 1993). Sensory profiling of a product can guide product development teams on what to change to match the consumer's desired sensory profile, to get closer to a benchmark, to detect detailed differences created by a change of an ingredient, etc.

The hypothesis that consumers are able to accurately describe products is more and more managed within the sensory science community. A first step in the development of effective techniques was the exploration of some methods like Repertory Grid Method, or the emergence of new ones as Sorting, Projective Mapping (known also as Napping®) or Flash Profiling. Several researches have already used these methods and focused on their validation with panels who have received different levels of training (Faye et al., 2004; Nestrud and Lawless, 2008; Perrin et al., 2008) but not much was done to assess the comparative applicability of all this methods with the use of naïve consumers panels.

The sorting task aims to detect meaningful sensory characteristics within pairs of samples that explain similarities and dissimilarities within the investigated sample set. The method was applied to various sorts of products: breakfast cereals (Cartier et al., 2006), plastic pieces (Faye et al., 2004) and beers (Chollet and Valentin, 2001) to mention a few. It consists of sorting products into groups according to their similarities. The method has the advantage that it can be applied to a large sample set but it often needs to be completed by a verbalization task in order to describe the groups formed and to explain the dimensions of the resulting perceptual map (Popper and Heymann, 1996).

Projective Mapping, and its variant Napping®, are profiling methods that were developed (Risvik et al., 1994, Pagès, 2005) in order to collect an Euclidian configuration for each assessor in a single sensory session. Samples, simultaneously presented, are positioned by each assessor on a tablecloth or a blank paper according to the differences/similarities (sensory distances) present between them in such a way that the smaller the distance separating two samples, the more similar they are (Perrin et al., 2008). The positioning criteria and their importance are chosen on an individual basis by each assessor, which makes Projective Mapping a flexible and spontaneous procedure.

Data are entered as position coordinates (x and y, with an origin that can be placed anywhere (Perrin et al., 2008) and the judgments of the assessors are equally taken into account. However, the number of samples presented should be limited to sets of 10-20 samples in order to limit fatigue or adaptation (Schifferstein, 1996). Similarly to Sorting, Projective Mapping does not describe the product itself and needs to be completed with either instrumental or sensory data (Pagès, 2005) or with a verbalization task to better understand the perceptual dimensions.

Flash profiling was defined by Siefferman (2000, 2002) as a combination of Free Choice Profiling with a comparative evaluation of the product set. It is a flexible method meant to position products rapidly according to their sensory attributes. It proved to be as satisfactory as conventional profiling when products are very different in terms of sensory attributes (Dairou and Sieffermann, 2002). However, when the tested products belong to the same product category or to similar product categories, Flash Profiling appears to be more discriminating than conventional profiling (Delarue and Sieffermann, 2004).

The repertory grid method (RGM) is based on the theory of personal construct psychology developed (Kelly, 1955). It associates meanings with products as bipolar constructs and results in a broad picture of how decisions are taken (Russell and Cox, 2003). For an example an assessor can be given three drinks and he/she may say that two of them are fruit-based while the third one is a dairy drink. 'fruit-based drink" is a construct in this context. In general, RGM is conducted in two sessions. The first one is dedicated to the attribute generation where products are presented in triads to the assessors who are asked to differentiate 2 samples from a third within each triad and explain why. The second session is a rating session in which samples are given scores for each of the elicited attributes. Assessors can also be asked to define a scale to quantify each perceived construct (attribute). This way, each assessor builds his/her own attributes and scales which are then used as in Free-Choice Profiling (FCP), in order to obtain a configuration of N objects in K dimensions (Williams and Langron, 1984).

The objectives of this study were, (a) to prove whether naïve consumers are able to describe hot beverages and generate relevant attributes by four descriptive methods: sorting, projective mapping, Flash Profiling and Repertory Grid Method; (b) to compare the consumers' description of the same sample set in 2 countries of the EU: the United Kingdom and France, looking at the influence of the language in the description; (c) to critically compare the applicability of the four methods and to correlate the outcomes to a trained consumer panel description via quantitative descriptive analysis.

2. Material and Methods

2.1 Sample set

A sample set of 8 hot beverages (seven samples plus one of them repeated), was used to perform the four descriptive methods. In the quantitative description by the trained panel (QDA), all of the 7 samples were evaluated by duplicate. The 7 products were selected in order to cover a wide flavour space, with distinctive sensory properties.

2.2. Sample preparation and serving designs

The drinks were served warm (at 70-75 ºC) immediately after preparation, in 3-digit coded paper cups. Tasting evaluation was performed in individual booths, under white light and at room temperature. Samples for Sorting, Projective Mapping and Flash Profiling were delivered to consumers in the three cases all at once, to be compared. In session 1 of the Repertory Grid Method (RGM) samples were presented to consumers in 3 triads were the samples were rotated, to avoid position and carry over effects, using a presentation design following a MOLS design (multiple orthogonal Latin squares). For the second session of the RGM and the quantitative descriptive analysis, samples were presented sequentially, following a Latin square design.

2.3. Panels

2.3.1. Trained panel

A panel of 11 trained assessors tasted and described the same sample set as the consumers did. Panellists were trained in the assessment of the category of products, varying in tasting experience from 1 to 15 years.

2.3.2. Consumer panels

Sorting, Napping, Flash Profiling and Repertory Grid methods were tested using a different panel of 24 naives consumers each, who were recruited by a recruiting agency according to

the following screening criteria: frequent consumers of the category in study (hot beverages), not rejecters of milk or sugar, ages between 18 and 65, 50% males, 50% females.
At the end of each tasting session, the consumers were asked to fill in a feedback form and answer questions related to the understanding, ease and time-effectiveness of each profiling method they used.

2.4. Profiling methods

2.4.1. QDA

Samples were completely rotated and 2 repetitions were completed. The evaluation proceeded in 3 sessions of 2 hours each:
Session 1 – Training: *S*amples were presented to the panellists in pairs. A list of 42 attributes corresponding to the hot beverage product category was used for the assessment. In this step the panellists rated the pair of samples perceived intensities on 150 mm closed-end unstructured scales. This task informed about whether an attribute was perceived by the panellists, and allowed assessing the degree of consensus in the ratings. If discrepancies in attributes or ratings were detected, an open discussion was prompted, in order to arrive to a consensus. Attributes selected for the data collection step were the ones utilized by at least half of the panel (23 attributes in total).
Sessions 2 & 3 – Sensory evaluation & repetition: *S*amples were presented to the panellists in a sequential monadic way following a Latin Square Design generated by Fizz (FIZZ 2.40B, Biosystems, France). They entered their intensity ratings by logging in a FIZZ QDA session built for 7 products and including the 23 selected attributes, using closed-end unstructured 150mm scales displayed on computer screens. *Session 3*: was a repetition of session 2 in order to check the performance of the panel as well as the reproducibility and attribute interactions as used by the panellists.

2.4.2.Sorting

Sorting was performed in one session of approximately 40 minutes, including briefing. The 8 samples were given all at once and consumers were then asked to observe, smell and taste them and then to group them according to their similarities, the number of groups formed should be no less than 2 and no greater than 7. Similar procedure was applied by Cartier et al. (2006), Faye et al. (2004), Chollet and Valentin (2001), and Giboreau et al. (2001). Once the sorting was performed, the consumers were asked to describe each of the groups they formed by giving their grouping criteria using sensory descriptors (Cartier et al., 2006). They reported their answers on individual ballots.

2.4.3. Projective mapping

The exercise was completed in one session of about 40 minutes, including briefing. Samples were presented simultaneously and the consumers were asked to observe them, smell them and taste them and then to position them according to individual criteria on a blank A3 sheet according to perceived similarities and/or differences in such a way that: (a) two samples are close to each other if they're similar; (b) two samples are far from each other if they are different (Pagès, 2005).

After positioning the cups on the A3 sheet, the consumers were asked to describe the samples and/or groups of samples (Perrin et al., 2008). The description comments were written on the A3 sheet next to each cup and/or group of cups. The cups were left on the sheets and their (x,y) coordinates were measured by the test leader and entered on a Excel sheet.

2.4.4. Flash Profile

FP was carried out in one session of one hour, including briefing. Coded samples were presented simultaneously and the consumers were asked to observe, smell and taste them in order to generate descriptors (on an individual basis). The next step was to rank all samples from ''least'' to ''most'' according to each attribute (Delarue and Sieffermann, 2004; Dairou and Sieffermann, 2002).

The flash Profiling task was performed by a consumer panel in Banbury, UK and another one in Paris, France.

2.4.5. Repertory Grid Method

RGM was conducted in two sessions of an overall duration of 2,5h including a 20 minute break, and a briefing part.

Session 1: Attribute generation: Samples were presented in triads following a balanced rotation in such a way that all samples appear at the same frequency. Each consumer was given 3 trays in order to allow the tasting of the whole the sample set. For each tray, the consumer was asked to pick the odd sample out and to explain in which way it is different from the two other samples and also in what way these two samples are similar (Russel and Cox, 2003; Monteleone et al., 1997; Thomson and McEwan., 1988, Gonzales-Thomas and Costell, 2006). The consumers entered their answers on individual ballots and had a break of 20 minutes after finishing the attribute generation for the 3 trays. During this break, the test leader entered the generated attributes on intensity rating ballot sheets to be used in the second session. Synonyms as well as attributes related to liking were omitted.

Session 2: Intensity rating: Samples were presented in a sequential monadic series. The task consisted in rating the perceived intensity of the attributes generated in the first session for each sample, using 150mm closed-end, unstructured scales with the extremes "Not at all" and "Extremely". Thus, each assessor had his/her own list of attributes.

2.5. Data analyses

Depending on the profiling method used in this project, the attribute elicitation can be combined or not with a rating and/or ranking task and thus the type of generated data is either qualitative, quantitative or both. This implies a cautious selection of the most adapted multivariate analysis. For both Sorting and Projective Mapping, frequency of mention scores of synonyms and repeated sensory attributes were combined and considered as one variable in the data analysis (i.e.: dull, mild, bland would be entered as dull/mild/bland with a frequency of mention that is the sum of the frequencies of each word).

Table 1 explains the choice of the data analysis method for each profiling technique used to obtain a sensory map.

In the case of the PCA, GPA and MFA methods, the number of relevant dimensions was selected by looking at the scree plots; the number of factors to be kept corresponding to the first turning point found on the curve. For MDS analysis the Shepard diagram was used to observe any ruptures in the ordination of the distances which helped choosing the number of dimensions.

Apart from these methods, and for all the profiling techniques, Hierarchical Cluster Analysis was applied in order to group samples as per their complete sensory profiles. More detailed description of the methods of choice for each technique can be found below.

The following software programs were used for data processing: Senpaq 4.2 (QIstatistics, UK) for checking trained panel performance, Fizz Acquisition and Fizz calculations (Fizz 2.40B, Biosystemes, France) for the trained panel data collection and analysis, and XLSTAT 2008.1.03 (Addingsoft, USA) for the rest of the data analysis.

2.5.1. Sorting

Individual similarity matrices [Products X Products] were built for each consumer by entering the number of times at which each pair of samples was sorted in the same group.

The sum of the similarity matrices, of all consumers, resulted in an overall similarity matrix which was then transformed into a dissimilarity matrix by subtracting the matrix data from the total number of consumers. The overall dissimilarity matrix was processed by XlStat functions (Describing Data and Multi-Dimensional Scaling (MDS)) to deliver a proximity matrix and further on a configuration of the sample set in two dimensions, dim1 and dim2. As the defined measure of dissimilarity between samples can not be considered as numerical data, a non-metric; Ordinal 1 model was applied. In order to explain these two dimensions, a matrix (Products x Attributes) was built by listing all the attributes generated by the consumers and by summing their frequencies of mention for each product. Spearman correlation was then determined between the attribute frequencies of mention and the product positions in dim1 and dim2. This way it was possible to plot the attributes over the product configuration obtained from the MDS calculations.

2.5.2. Projective Mapping

Multiple Factor Analysis (MFA) was performed with XlStat on data obtained from the positioning of the products on the A3 sheets as well as the attributes each consumer generated to describe the products. The product coordinates were measure in centimetres and attributes listed together with their frequencies of mention across the consumer panel, resulting in three tables: x, y, attributes (Figure 1).

2.5.3. Flash Profiling and RGM

Individual matrices for each consumer (Products x Attributes) were built in order to enter product rankings (from FP) or intensity ratings (from RGM). A GPA was then performed on the 24 matrices in order to obtain the product and attribute configurations.

2.5.4. Quantitative Descriptive Analysis

Summary statistics involving the calculation of means, standard deviation and ranges were carried out aiming to get an overview of the complete data set.

Panel performance was evaluated with the use of Senpaq 4.2 (QIstatistics, UK) package, performing ANOVA with a fixed model, checking discriminatory ability, reproducibility and scale usage of each panellist, as well as each panellist's contribution to interaction.

Analysis of variance (ANOVA) with the use of a mixed model (assessors as random effect) was performed for testing product set significant differences. Multiple comparisons were preformed with Tukey's test at 5%. A principal component analysis (PCA), based on the Pearson's correlation matrix, was conducted on the means of the attributes presenting significant differences. PCA allows the profile data to be summarised in a smaller number of dimensions than the total attributes in the profile (principal components). Each component represents a certain percentage of the total information or variability of the original profile data. Samples and sensory attributes can be projected onto these components and summary plots can be produced. PCA is a statistical tool that helps to summarise and therefore, communicates better the results from descriptive panel profiling.

2.5.5. Other Statistical Analyses

Hierarchical Clustering Analysis (HCA) was performed in order to highlight product clusters as perceived by the trained panel and by the consumer panels. Method details were: dissimilarity: Euclidean distance, agglomeration method: Ward's method, automatic truncation.

Consumer cross-method comparisons were based on visual similarities of the product positioning on the profiling maps, their description, the identification of the internal repetition, as well as the number of sensory attributes generated.

Maps were correlated with the trained panel map using a MFA in such a way that mean scores from the trained panel were considered as an additional data set to consumers' individual data

3. Results and discussion

3.1. Quantitative Descriptive Analysis – trained assessors

Taking into consideration the ANOVA results obtained through the PCA analysis of the descriptive data conducted by the trained panel, revealed a clear discrimination of the sample set following 18 attributes related to aftertaste, flavour and mouthfeel. The first 2 dimensions explained over 84% of the variability, describing the most important sensory attributes of this sample set, main perceptual directions are represented in the map. Hierarchical Clustering Analysis (HCA) revealed 3 groups of products (Figure 2).

3.2. Sorting – naïve consumers

When performing a MDS, the difference is measured through the stress, several variations of which have been proposed: Raw, Normalized, Kruskal's 1&2. In this work, the best fit delivered by the MDS was obtained at a Kruskal's stress (1) K=0.089. However the panel's performance was not satisfactory as the internal repetition was not mapped in a relevant way (P3 and P3 Rep are far from each other). After manually overlaying the product and attribute maps, it was observed that the products were well described by the attributes (figure 3). For instance: P1 was described as weak in taste, smell and aroma, mild and with a vanilla flavour, P2 was described as dark and rich in colour with a strong taste, P3 was described as creamy, sweet and not bitter and P4 as being bitter, sharp and intense. However, P1 occupied an intermediate position between P2 (a bitter and strong sample) and P4 (an intense and unsweetened sample) while it was expected to be placed far from both samples according to the QDA results.

The low performance of the sorting task could be explained by the fact that the sensory assessment as described by this method tends to group the products following one sensory dimension only: consumers sorted the samples following one main criterion (appearance, smell or taste) even if many grouping options could exist for each product; that could mean that the final configuration might not contain enough sensory attributes for a multidimensional map to be relevantly built. Adding a verbalization task to sorting allowed a better understanding of the obtained results. The generated attributes were in fact consistent with each sample's sensory profile. For instance, even if P1 (a bland, milky and sweet sample) was placed between P2 (intense, bitter and sweet) and P4 (unsweetened, intense), it was described as 'milky', 'bold', 'tasteless'.

Similar findings were obtained for breakfast cereals by Cartier et al. (2006) who compared sorting with an untrained panel to QDA. Moreover and when repeated, sorting results (product mapping and description) showed to be consistent over time (Cartier et al., 2006; Lawless and Glatter, 1990). Faye et al. (2004) on the other hand found that sorting of leather pieces with consumers gave results which were consistent with the trained panel's results in terms of product positioning and description. This can be explained by the complexity of the sensory profile of the investigated product and the level of the training.

The success of a sorting task with consumers would then very much rely on the differences within the sample set. If they are big, and involving multivariate perceptions (appearance, aroma, flavour, manual and oral texture, aftertaste), consumers would be indeed able to discriminate between them, and most probably would use similar attributes to describe those differences, resulting in a good consensus map of the samples. In the present study, most of the differences were in flavour and aftertaste, being all liquid products; these facts added up to the uni-dimensionality of the sorting task, explain the low performance of it as a mapping procedure. The hypothesis is that, in a more complex sample, consumers would naturally tend to use multiple sensory attributes and different dimensions in the sorting task (texture, flavour, aroma), in contrast to a rather plain food as a hot beverage is. More research would be needed to validate this hypothesis.

3.3. Projective Mapping – naïf consumers

Figure 4 shows the product positioning coming from the MFA of the projective mapping data. The map obtained was quite comparable to the QDA map, though the first two dimensions only explained 45% of the variability. Nonetheless, and apart from P1 which was placed on its own following the 'Milky/Creamy' vector, P2, P3, P3 Rep and P7 occupied the same space. These products were identified as being part of the same cluster in the QDA results of the trained panel. This means that the consumers were also able to identify these samples as similar to each other and different from the other products even if they could not perceive well the underlying differences present within the cluster (figure 4 shows small distances between these samples). However, even if the sample P3 and the internal repetition are closer to each other with the Projective Mapping results than they are with the Sorting results, the internal repetition did not seem to be obvious to the consumers. Apart from that, the discrimination of the samples was lower than expected; samples were not separated as per the clusters identified with the trained assessors.

Attribute wise, the basic attribute directions describing the sample set were identified and were aligned with the QDA descriptors; P2, P3/P3 Rep and P7 are driven by sweetness in general but also by coffee intensity and bitterness which they shared with samples P4, P5, P6, and P1 by milkiness/creaminess. The relevant positioning of the products showed that PM performed better than Sorting and this can be explained by the fact that PM is a bi-dimensional procedure where samples are placed on a surface following the perceived similarities/dissimilarities between them. Such task could be appropriate to highlight big differences but was not precise enough in pointing out more accurate differences between samples in the present study. Perrin et al. (2008) applied a similar approach where Napping® was combined to Free Choice Profiling and performed with both a trained panel and a panel made of wine professionals. The study was however more focused on the validation of Napping as a relevant profiling method and was not performed with naive consumers as it is the case here. The results obtained by Perrin et al. showed in fact that Napping gave a good product positioning with a rough description which was not as accurate as the QDA profile's description. Nestrud and Lawless (2008) however, stated that Projective Mapping would perform better than sorting as it provides richer data sets per assessor in addition to the fact that product positions on the 'nappe' can inform about their similarities.

3.4. Flash Profile – naïf consumers

3.4.1 UK Panel

The relative configuration of the samples obtained from the FP procedure is very similar to the one obtained from the QDA. As shown in figure 5, the consumers were able to perceive the differences and similarities between the samples with the first two dimensions explaining more than 82% of the variability. This is made even clearer when looking into the main attribute directions highlighted in the map, where it can be observed that 'bitterness' and 'colour intensity/coffee intensity' described P2, P4, P5 and P6. These attributes were negatively correlated with 'sweetness', 'smooth/creamy' and 'bland/weak' which described P1, P3/P3 Rep and P7. The 'Creamy' attribute seemed however not to be well understood or agreed on by the consumers as it covered a wide space (shaded area in figure 5). It can be explained by the fact that consumers do not have the same understanding of 'creamy' (and related words e.g.: cream, creaminess) and that creaminess is an integrated attribute associated to flavour, texture and pleasantness of food products (Tournier et al., 2007).

The product clustering obtained with the HCA (ellipses shown in figure 5) is the same as the one obtained from the QDA results. The panel's performance was highlighted by the fact that samples P3 and P3 Rep were positioned close to each other and both belonged to the same cluster.

These findings are in agreement with those from Dairou and Sieffermann (2002) who conducted a study comparing conventional profiling and flash profiling performed on 14 jams using respectively a trained panel and a panel made of students with a relevant experience in sensory methodology. The results obtained revealed that both methods produced similar results in terms of positioning and description of the jams, with flash profiling being faster but less self-explanatory than conventional profiling. Another study comparing the evaluation of dairy products (strawberry yogurts and apricot-flavoured fresh cheese) with Flash Profile and a conventional profiling (as QDA or Spectrum), both with a trained panel, showed that flash profiling was more discriminating than conventional profiling but that it was less adapted to the description of the investigated product category (Delarue and Sieffermann, 2004). Sieffermann and col. successfully proved that FP is a valid method for a quick product profiling with the use of trained assessors but did not explore the validity of its use with naive consumers as it is the case here; the present study gives a further step in proving the applicability of such technique as a profiling method .

3.4.2. French Panel

The French panel was also able to position the samples in a relevant way in comparison with the QDA product plot (figure 2), as well as to represent the expected differences and similarities present within the sample set. The first two dimensions explained about 84% of the variability (figure 6).

With regards to the attributes (figure 6), the main directions were similar to those observed for the outcomes of the QDA with the trained panel. The HCA also revealed the same product clustering as in the QDA results. The French panel was consistent in describing the differences/similarities within the sample set (P3 and P3 Rep came out close to each other on the GPA product configuration).

When comparing the FP results from the UK and French panels, it appears that both delivered similar product positioning, clustering and description (attribute directions), with the first two dimensions explaining more than 82% of the variability. Moreover, and as long as the 'creamy' attribute is concerned, it seems that the French panel, just like the UK panel, used this attribute to describe different perceptions. A translation of the French terms is given in table 2.

Apart from the main attribute directions shown in figure 6, the attribute correlation as given by the GPA also reveals some interesting well correlated attributes such as 'viande' (meat) , ‘odeur céréales’ (cereal smell), ‘gout aspartame’ (aspartame flavour) and ‘acidité’ (sourness).

Consumers from France and the UK are quite different in consumption habits and preferences for this category of products. The fact of finding very similar outcomes in both cases shows that Flash Profiling could be used regardless the liking patterns of the consumers, providing they are not rejecters of the category. Another potential application coming from the results of the present study is that FP could be applied in different languages to get more insight about consumer relevant attributes in the target country, and to compare to the attributes used by the trained panel.

3.5. Repertory Grid Method

The GPA results from the RGM data sets shown in figure 7 revealed a relevant relative positioning of the samples in the perceptual space, and also a pertinent description of the perceived differences; 83% of the variability was explained by the first two dimensions only. Both configurations (products and attributes) were consistent with the QDA results. The HCA product clustering results were identical to those obtained for QDA as well as for Flash Profiling and showed that the sample set could be discriminated into three clusters.

The comments made about the ‘creamy’ attribute as mentioned by the consumers in the FP task, can also be made for the RGM results. ‘Creamy’ is obviously a complex attribute that probably covers creamy mouthfeel, creamy flavour, creamy smell and creamy appearance. In addition to the main attribute directions, the attribute map obtained from the GPA revealed the following most correlated attributes: ‘biscuity’ scent, nut taste, cocoa taste, caramel, dry, sour, palatable, smooth, artificial, tangy taste and coffee strength.

RGM has already been performed with consumers but with the objective to understand consumers’ perception of a product rather than comparing consumers’ profiling with a trained panel’s profiling and the aim of obtaining a product perceptual map (Russel and Cox, 2004, applied to the description of meat products; Hersleth et al., 2005, applied to bread). Nonetheless, the results of these studies showed an alignment between the product positioning and description given by the consumers and those revealed by the trained panel.

3.6. Generated vocabulary

The four investigated product profiling methods proved to generate a fair amount of sensory relevant attributes (Table 3). As presented in the mapping results, there was consensus in the use of these attributes by the consumer panels and the core sensory attributes of the sample set matched the trained panel’s description. Table 3 gives the main results through the four profiling methods. It can be observed that even if Sorting and Projective Mapping were not very well suited for mapping objectives in this study, they proved to be an acceptable way of vocabulary development; they generated a high amount of attributes, although only 12% and 14% of these attributes were respectively relevant, after taking out non sensory words and grouping synonyms . These findings are very important as a way to validate the trained panel vocabulary, ensuring that what is analytically measured reflects relevant perceptions for consumers. It also proves that the methods tested in the present work could be used to generate vocabulary related to the product category in different target countries, to study which words are used by consumers to speak about the different aspects of a product sensory design.

It is worth mentioning that for all methods, hedonic and non-sensory attributes such as 'natural appearance', 'bad flavour' or 'nice smell' were not included in the data analysis. Others such as 'right sugar balance' or 'correct coffee balance' were considered respectively as 'sugar' and 'coffee', whereas attributes like 'not bitter' or 'no coffee flavour' were used as such.

3.6.1 Sorting

The top 20 most mentioned generated attributes are listed in table 4. They count for 20% of the generated vocabulary with over 53% of the frequencies of mention.

3.6.2. Projective Mapping

The most frequently mentioned sensory attributes generated are listed in table 5 and it can be seen that this table is quite similar to the one obtained with the sorting task. However, about 50% of the attributes generated by Projective Mapping were mentioned only once across the panel.

Projective mapping also generated 32 non-sensory words among which we can find some interesting ones such as: 'forgettable', 'has heart and soul', 'pleasant' and 'little to distinguish'.

3.6.3. Flash Profile

The vocabularies generated by the UK and French panels were quite similar and (table 6), 'sweet', 'bitter', 'colour', and coffee taste/strength were in the top 5 most mentioned words for both panels. Moreover, the top 20 most frequently mentioned attributes by both panels accounted for 73% and 75% of the attributes generated by the UK and the French panels respectively.

Table 6 also highlights the common attributes between the British and French panels (in bold). These account for more than 60% of the attributes generated if we exclude the synonyms (e.g. darkness=colour dark, smell = strength of smell, etc.). The UK panel also generated 3 non-sensory words: choice, 'nice horrible' and 'perfect recipe' whereas the French panel did not generate any.

3.6.4. RGM

The top 20 most frequently generated attributes with RGM are very comparable to those generated by FP (table 7). They represent about 40% of the generated attributes with 80% of the frequencies of mention. 'Sweet' was mentioned by nearly every consumer and 'colour', strength of taste', 'bitter' and by more than half of the panel.

The vocabulary generated by RGM and considered for the GPA analysis did not contain any non-sensory descriptors. This is most probably due to the fact that the vocabulary was elicited thanks to straightforward questions in the 1st session of RGM (task question : "Pick the odd sample out within the triad, explain in what way it is different from the two other samples and in what way these two are similar). Additionally, the first step in vocabulary generation for RGM (triadic elicitation) is screened down by the panel leader for allowing the rating stage of the sensory descriptors in the second step.

3.7. Correlation of consumer and trained panel results

FP and RGM were identified as the best suited methods in terms of product consumer profiling, accomplishing both mapping and description objectives. Their perceptual maps were subsequently correlated with the maps obtained from the QDA using MFA as a way to better understand the correlation between consumers' product attributes and the attributes from the trained panel. MFA was performed considering the mean scores from the trained panel as an additional data set to the consumers' individual data, this way the structure of the product configuration was maintained as per the consumer's description. MFA makes it possible to analyze several tables of variables simultaneously, and to obtain charts that allow studying the relationship between the observations, the variables and tables graphically. The originality of this method is that it allows visualizing in a two or three dimensional space, the tables, the variables, the principal axes of the analyses, and the individuals. In addition, the impact of the other tables on an observation can be studied by simultaneously visualizing the observation described by the all the variables and the projected observations described by the variables of only one table (Escofier and Pagès, 1984).

The correlation map via MFA showed that the two first dimensions explained up to 65% of the variability for FP, with a high correlation (Rv=0.91) and 60% for RGM (Rv=0.88). Moreover, the descriptors generated by the consumers in both cases reflected well those used by the trained panel. For instance, the 'Sweetness' perceptual direction as described by the trained panel, is positively correlated with descriptors such as 'sweet', 'sweetened' and sugar' which were generated by the consumers, the 'Bitterness' direction as described by the trained panel explains consumer words such as ‘strong taste’, ‘bitter’ or ‘intense’. These results confirm the suitability of these two techniques as a means of obtaining a consumer relevant product mapping as well as their ability to generate a list of words to describe the samples in “consumers’ language”.

Nevertheless, as mentioned before, the consumers’ perception of ‘creamy/creaminess’ was not well correlated to the trained panel assessment. Moreover, the QDA results revealed a more detailed description of the samples in some aspects, not described in depth by the consumer panels, particularly on mouthfeel perception and aftertaste/afterfeel. These findings suggest that there is a limit in the quality of information that can be obtained via consumers’ description. The use of MFA to correlate data sets from different profiling methods was particularly useful in this study as it provided a graphical idea of how the data sets related, i.e. whether the attributes were explained in the same way by both panels being compared (trained panel and consumer panel), and what the discrepancies were (see figure 8 as an example of the correlation of RGM and QDA results). This allowed the observation that the “creamy” perception was not quite well described by consumers.

3.8. Practical considerations and applications

A straightforward application of the evaluated consumer profiling techniques is when there are time constraints, when a trained panel is not available or a new category of products is being tested in which the internal panel is not well trained.

Nevertheless, it has to be pointed out that the objective of this study was not to find new methods to replace descriptive techniques with trained panels but to further understand how methods as QDA relate to consumer perception.

Moreover, trained panel descriptive measurements perform better in various cases: when there is a need to compare samples in different moments in time and also when comparing different sample sets with few samples in common, as although not absolute, when the panel is well calibrated and maintained, trained panel measurements are stable both in time and within a certain perceptual space. Another example is when small differences have to be described; or as aforementioned, when the difference lays in more complex perceptions like mouthfeel and afterfeel, in those cases a panel with intensive training is still needed, and could be a crucial source of information in product development applications.

The other objective of this research was to critically compare the effectiveness and application of different profiling methods, for the first time done with multiple techniques, on the same sample set, and with different panels of naïf consumers, allowing a thorough comparison. The results showed that Sorting, Projective Mapping, Flash Profiling and Repertory Grid Method, all could be applied as a means of generating vocabulary in hot beverage samples; however, FP and RGM present the advantage of being also well suited as descriptive techniques for a quantitative product mapping. However, one limitation has to be highlighted for FP; being a comparative method, the number of samples that can be assessed by FP is limited (and would depend on the category), while in RGM and QDA, as samples are presented in a monadic sequence, there is no potential limitation regarding maximum number of samples.

FP and RGM data can be used with confidence to correlate with sensory data from a trained panel, or also with physical or chemistry data. Furthermore, the correlation of FP or RGM and QDA would be a strong validation of the descriptive panel glossary. The feasibility of quantifying consumer description in a relevant way via FP or RGM also opens the door of the exploration of consumers' vocabulary with a more statistical validity (contrary to traditional qualitative exploration methods as focus groups). As an example, some useful questions could be answered: "which consumer attributes discriminate better between two samples"?, "what attributes are more relevant to the target consumers"?, "do consumers have the same understanding of a particular attribute?, "what different words describe the same sensory stimuli for a naive consumer?", resulting in an interesting guidance to product developers. This could also be a way of identifying important attributes to consumers with the objective of having them into account in further quantitative affective testing of the category (central location test or home use test), particularly when the target country has a different language than the one managed by the research team.

4. Conclusions

The results obtained in this study proved that consumers were able to generate relevant attributes and to quantify differences between hot beverage samples.

Sorting and Projective Mapping performed poorly in terms of product discrimination and repeatability, but both are quick and user-friendly techniques that could be used when a broad and rough description is needed.

Flash Profiling and RGM were accurate in terms of mapping and clustering; these methods also produced a relevant and rich description. Consumers' measurement by FP and RGM (understood as product positioning and description) was very comparable to the measurement performed by the trained panel, however, when describing mouthfeel (e.g. creaminess) the consumers' description was not enough detailed or consensual.

The good performance of FP and RGM with consumers makes them reliable methods which can be recommended for obtaining a description when a quick profiling is needed or when the panel of trained assessors is not available or not trained in that particular product.

Sorting and Projective Mapping on the other hand can be used as easy and quick product profiling procedures when an accurate profiling is not needed and when a rough description is sufficient.

All the methods studied could be used as direct feedback from target consumers, as good vocabulary development tools; however FP and RGM being more accurate at profiling can be used also to correlate to data from a trained panel and even to physical and chemistry data.

In this work, trained panel attributes reflected well consumers' perception of the product category in study, serving as a validation of the trained panel glossary. In general, trained panel descriptive measurements still would perform better in some applications and can not be substituted, particularly when complex attributes are involved, as in this case mouthfeel and afterfeel, or small differences need to be characterized.

The profiling techniques researched here are seen as complementary to the information given by a trained panel, being very powerful tools to obtain direct feedback from consumers. Next steps would be to study "how far can these methods go", i.e. are they applicable to most product categories? In which cases are they better suited or not suited at all?

Acknowledgments

The authors would like to thank Marie-Christine Marcuz from Kraft Foods LU-Biscuits (Paris, France), Catriona Lakemond from Wageningen University (NL) and the RD&Q team at Kraft Foods Banbury (UK).

References

ASTM (1992). Quantitative Descriptive Analysis (QDA). ASTM Digital library. DOI: 10.1520/MNL10523M

Cartier, R., Rytz, A., Lecomte,A., Poblete, F., Krystlik, J., Belin, E., Martin, N. (2006). Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. Food Quality and Preference 17, 562–571.

Chollet, S., Valentin, D. (2001). Impact of training on beer flavour perception and description: are trained and untrained panellists really different? Journal of Sensory Studies, 16, 601–618.

Dairou, V.; Sieffermann, J.-M. (2002). A comparison of 14 jams characterized by conventional profile and a quick original method, the flash profile. Journal of Food Science, 67(2), 826-834.

Delarue, J. ; Sieffermann, J.-M. (2004). Sensory mapping using Flash profile. Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. Food Quality and Preference, Volume 15, Issue 4, 383-392.

Dijksterhuis, G. (1997). Multivariate data analysis in sensory and consumer science. Food and Nutrition Press, Inc. pp. 15. ISBN: 0-917678-41-9

Escofier, B. & Pagès, J. (1984), L'analyse factorielle multiple: une méthode de comparaison de groupes de variables, Data analysis and informatics, III, Diday E. edn, Elsevier Science, Amsterdam, pp. 41–56

Faye, P., Brémaud, D., Durand Daubin, M., Courcoux, P., Giboreau, A., Nicod, H. (2004). Perceptive free sorting and verbalization tasks with naive subjects: an alternative to descriptive mappings. Food Quality and Preference 15, 781-791.

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 descriptive method of tactile properties. Food Quality and Preference, 12, 311–322.

Gonzales-Tomas, L. & Costell, E. (2006). Sensory evaluation of vanilla-dairy desserts by repertory grid method and free choice profile. Journal of Sensory Studies, 21, 20-33.

Hersleth, M., Berggren, R., Westad, F. & Martens, M. (2005). Perceptions of bread: a comparison of consumers and trained assessors. Journal of Food Science, 70 (2), S95-S101.

Kelly, G. A. (1955). The psychology of personal construct: A theory of personality. New York: Norton.

Lawless, H.T. & Glatter, S. (1990). Consistency of multidimensional scaling models derived from odor sorting. Journal of Sensory Studies, 5, 217-230.

McEwan, J.A., Thomson, D.M.H. (1988). An investigation of the factors influencing consumer acceptance of chocolate confectionery using the repertory grid method. In Food Acceptability, D.M.H. (Ed.). Elsevier Applied Science, London, pp. 347-362.

Monteleone, E., Raats, M.M., Mela, D.J. (1997). Perceptions of starchy dishes: application of the repertory grid method. Appetite, 28, 255-265.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: application to the study of ten white wines from the Loire Valley. Food Quality and Preference, 16(7), 642-649.

Perrin, L., Symoneaux, R., Maître, I, Asselin, C., Jourjon, F., Pagès, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: Case of ten wines from Loire Valley. Food Quality and Preference, 19; 1-11.

Popper, R., & Heymann, H. (1996). Analyzing differences among products and panellists by multidimensional scaling. In Naes, T. and Risvik, E. (Eds.), Multivariate analysis of data in sensory science, pp. 159–184. Amsterdam: Elsevier.

Risvik, E., McEvan, J.A., Colwill, J.S., Rogers, R. and Lyon, D.H. (1994). Projective mapping: A tool for sensory analysis and consumer research. Food Quality and Preference 5, 263–269.

Russel, C.G. & Cox, D.N. (2004). Understanding middle-aged consumers' perception of meat using repertory grid methodology. Food Quality and Preference 15, 317-329.

Schifferstein, H.N.J. (1996). Cognitive factors affecting taste intensity judgments. Food quality and Preference, Second Rose-Marie Pangborn Memorial Symposium, 7(3-4), 167-175.

Sieffermann, J.-M. (2000). Le profil flash - un outil rapide et innovant d'évaluation sensorielle descriptive. In AGORAL (2000). XIIèmes rencontres ''L'innovation: de l'idée au succès'' (pp. 335–340), Montpellier, France.

Stone, H. and Sidel, J.L. (1993). Sensory Evaluation Practices. Academic Press, Inc. pp. 143-270. ISBN 0-12-672482-2.

Thomson, D.M.H. & McEwan, J.A. (1988). An application of the Repertory Grid Method to investigate consumer perceptions of foods. Appetite, 10, 181-193.

Tournier, C., Martin, C., Guichard, E., Issanchou, S., Sulmont-Rossé, C. (2007). Contribution to the understanding of consumers' creaminess concept : A sensory and a verbal approach. International Dairy Journal, 17, 555-564.

William, A.A. & Langron, S.P. (1984). The use of free-choice profiling for the evaluation of commercial ports. Journal of Science of Food and Agriculture, 35, 558-568.

FIGURE CAPTIONS

Figure 1: Schematic exemplifying view of how projective mapping data were treated: x1 and y1 are the coordinates of P1 for consumer 1, x2 and y2 are the coordinates of P1 for consumer 2, etc. Attributes were recorded with the frequency of mention per product; “sweet” was mentioned 6 times for P1, etc.

Figure 2: PCA product map from QDA assessment (attributes not shown, main attribute directions are represented in the map). Ellipses highlighting product clusters as revealed by HCA

Figure 3 : Product configuration obtained by MDS from the Sorting task (Overlaying of Spearman correlation of attributes)

Figure 4: Product configuration obtained by MFA from the Projective Mapping task. Vectors show the most correlated attribute directions

Figure 5: Product configuration obtained by GPA from the Flash Profiling with the UK panel. Vectors show the most correlated attribute directions. Shaded area shows the “creamy perceptual space”. Ellipses showing clusters from HCA

Figure 6: Product configuration obtained by GPA from the Flash Profiling with the French panel. Vectors show the most correlated attribute directions. Shaded area shows the “creamy perceptual space”. Ellipses showing clusters from HCA

Figure 7: Product configuration obtained by GPA from the RGM task. Vectors show the most correlated attribute directions. Shaded area between two vectors shows the “creamy perceptual space”. Ellipses showing clusters from HCA

Figure 8: Correlation of RGM and QDA results obtained by MFA. Vectors shoy the most correlated attribute directions generated by consumers. Shaded area between two vectors shows the “creamy perceptual space”. Names of QDA attributes are not shown.

TABLES

Table 1: Product profiling methods versus adapted multivariate analysis technique

Method	Data analysis	Reason
Sorting	MDS	- Maps the pattern of similarities or dissimilarities perceived among a set of products/objects by computing the frequencies of sorting which each pair of products in the same group
Napping®/ Projective Mapping	MFA	- Analyzes several tables of variables which differ in number and nature from one another - Within a table, the variables must be of the same nature (quantitative or qualitative) - Integrates different tables of variables describing the same observations
Flash Profiling **Repertory Grid Method**	GPA	- Reduces the scale usage effects - Delivers a consensus configuration (consensus in the use of attributes come from the usage of the same/similar attribute by different panellists) - Allows to compare the proximity between the terms that are used by different assessors to describe products
Quantitative Descriptive Analysis	PCA	- Delivers a consensus configuration (attributes are also consensual between panellists) - Allows to compare the proximity between the terms that are used by different assessors to describe products

Table 2: Translation of French discriminating attributes

French attributes	English translation
Amertume	Bitterness
Puissance gout	Taste strength
Puissance arome	Aroma strength
Intensité de la couleur	Colour intensity
Sucré	Sweet
Gout lacté	Milky taste
Crémeux	Creamy
Saveur creme	Creamy flavour
Gout caramel	Caramel taste
Acidité/saveur acide	Sour/sour flavour
Odeur céréales	Cereal smell
Café grillé	Roasted coffee
Viande	Meat
Gout aspartame	Aspartame taste
Onctuosité	Creamy mouthfeel

Table 3: Cross-method and Cross-country key figures for generated vocabulary

	Sorting	PM	FP UK	FP France	RGM
Number of attributes generated	233	247	136	118	162
Average per consumer	8.6	9.2	5.3	6.2	6.3
Number of attributes excluding repetitions and non-sensory words	27	34	54	45	49

Table 4: Top 20 most frequently mentioned attributes with Sorting

Attributes	Frequency of mention	Attributes	Frequency of mention
sugar/sweet	14	not sweet	4
bitter	13	watery	4
strong	12	bland	3
cream	9	medium coffee	3
dark	9	milk	3
weak	9	no coffee taste	3
mild	8	pale	3
smooth	7	slightly bitter	3
rich	5	strong smell	3
mellow	4	strong taste	3

Table 5: Top 20 most frequently mentioned attributes with Napping®

Attributes	Frequency of mention	Attributes	Frequency of mention
sweet	18	mellow	4
bitter	13	milk	4
creamy	13	pale colour	4
strong	12	coffee taste	3
rich	8	mild	3
weak	8	mild flavour	3
dark	7	no flavour	3
bland	6	not bitter	3
no coffee taste	5	not sweet	3
watery	5	nutty taste	3
aftertaste	4	smooth	3
dark colour	4		

Table 6: Top 20 (16 for the French) most frequently mentioned attributes with FP for the British and French panels. In bold are highlighted the common attributes between the two panels.

UK Panel		French Panel		
Attributes	**Frequency of mention**	**Attributes**	**English translation**	**Frequency of mention**
strength/strong/strongest	17	**saveur sucrée**	Sweet taste	18
sweetest	14	**amargor/amer**	Bitter	13
bitter	11	**couleur (intensité)**	Colour intensity	12
colour/appearance	10	**gout lacté/arome lait**	Milky taste/milk flavour	10
smooth/smoothness	7	gout café	Coffee taste	5
creaminess	5	**intensité d'odeur**	Smell intensity	5
aroma	4	odeur café	Coffee smell	4
darkness	4	acidité	Acid/sour	3
richness	4	arome café	Coffee aroma	3
smell	4	**couleur marron**	Brown colour	3
after taste	3	**gout (intensité globale)**	Taste (overall intensity)	3
taste	3	aqueux	Watery	2
bland	2	**arome (intensité globale)**	Aroma (overall intensity)	2
colour dark	2	**café**	Coffee	2
depth / strength of coffee	2	corps	Body	2
milkiness	2	**onctuosité**	Smoothness	2
nuttiness	2			
smell intensity/intensity	2			
smooth flavour/taste	2			
strength of aroma/smell	2			

Table 7: Top 20 most frequently mentioned attributes with RGM

Attributes	Frequency of mention	Attributes	Frequency of mention
sweet	24	flavour/flavoursome	3
colour	19	mild/bland	3
strength of taste	19	coffee smell	2
bitter	14	dry	2
creaminess	9	mellow	2
smooth	6	palatable	2
weak	6	rich	2
coffee taste	4	smell	2
milkiness	4	strength of smell	2
strength of aroma	4	thick	2