Projective mapping based on choice or preference: an affective approach to projective mapping.

Paula Varela, Ingunn Berget, Margrethe Hersleth, Mats Carlehög, Daniele Asioli and Tormod Næs*

Nofima AS, P.O. Box 210, 1431 Ås, Norway

*Corresponding author: Paula Varela

Telephone: +47 45426026

Fax: +47 64943314

E-mail: paula.varela.tomasco@nofima.no; pvarelatomasco@gmail.com

Abstract

This work explores a new affective approach to projective mapping, based on consumers' choices or preferences. Two sessions, one week apart, were performed with the same consumers, using whole bread as a case study. Overall liking ratings (OL) were gathered in blind conditions and samples were also profiled by a trained panel using generic descriptive analysis. Three projective mapping tests were performed in different scenarios. Consumers' categorization and product descriptions were explored when consumers based their positioning on the products' similarities and differences (analytical approach, "classic napping") both in blind and informed conditions, and when consumers were focusing on their preference or choice (affective approach). The affective approach to projective mapping successfully revealed consumers' drivers of liking and choice from a holistic perspective, where consumers summarized their main drivers for categorizing products as they would do when choosing in real life situations, based on their preferences.

Keywords: projective mapping; napping; affective; consumers; drivers; preference; choice.

1. Introduction

Projective mapping (also known as Napping®) followed by a descriptive step has been extensively used with consumers in the last years as an alternative tool for the description of products and packs. It is considered a holistic approach to product profiling, closer to what happens in a choice event when compared to classic descriptive or attribute-based techniques (Varela & Ares, 2012; Valentin et al., 2012). Built on the perception of similarities and differences, it encourages the generation of a global representation of the products, which is usually hindered when consumers are directly asked about multiple particular attributes. Holistic methods enable to identify the main attributes that account for the differences among the samples without forcing consumers to focus on specific characteristics (Ares & Varela, 2012). In addition, projective methods make it possible to capture more spontaneous responses than other, more directive, techniques (Guerrero et al., 2010). The projective mapping (PM) task can involve the perception of similarities and differences from an intrinsic (sensory) perspective, from an extrinsic (pack, labelling, etc.) perspective, or from both (Carrillo, Varela, & Fiszman, 2012a), generally considering product objective characteristics for categorization rather than liking as the main parameter. Despite this, consumers often use hedonics or benefit-related terms together with the product and pack descriptive characteristics. This can be used to relate product characteristics to marketable features and consumer preferences (Ares & Varela, 2012) and is an approach that has been applied successfully to explore sensory and non-sensory stimuli, such as the influence of packaging information – e.g. nutritional and health claims – on consumer perception (Carrillo et al., 2012a; Carrillo, Varela, & Fiszman, 2012b; Miraballes et al., 2014; Varela et al., 2014).

When optimizing food products, the general practice has been to ask consumers about liking; the sensory properties would be characterized in parallel by a trained panel, in a preference mapping type of exercise (van Kleef et al. 2006). However, trained assessors may describe the product differently, so sensory characterization based on consumers' direct input may have greater external validity (Ares & Varela, 2012). In this sense, overall liking (OL) has been gathered jointly with PM data in some studies in order to draw conclusions on drivers of liking (Ares et al, 2010; Torri et al., 2013) and to better understand the changes in hedonic response in different mapping scenarios (Carrillo et al., 2012b). In a study by Ares et al. (2011), after doing a PM with real samples of powdered orange juice consumers were asked about their ideal product to be mapped. The results were similar to those of external preference mapping. Withers at al. (2014) have used taxonomic sorting, a holistic method also based on sample categorization, to generate diagnostic sensory data directly from target consumers by external preference mapping. Generally, hedonic descriptions or OL have been considered as supplementary variables in PM data.

From a different perspective, King, Cliff & Hall (1998) compared PM to a "structured PM" to map snack bars, where they used labeled axes in the PM space: the x-axis was defined as "liking" (low - high) and the y-axis as "usage" (treat - meal replacement). They found the proposed method less discriminating than PM, but only 24 consumers participated in this study. To our knowledge, there have been no other approaches to PM from an affective perspective, with liking or preference explicitly driving sample categorization.

Consumers in affective tests act in an integrative fashion, basing themselves on global sensory and non-sensory stimulation from the product – in contrast to the analytical testing frame of mind in descriptive testing (Lawless & Heymann; 2010; Jaeger, 2006). More concretely, since consumers are integrated and organized wholes, as highlighted by Maslow (1954), in real buying and eating situations they take a certain number of attributes (sensory and non-sensory) into account when performing food choices or declaring their preference (Asioli et al., 2017). Thus, consumers would cognitively focus on products differently when describing as opposed to stating their preference or choice.

With this background, it is of great interest to study how consumers approach the PM task when preference or choice is used as a criterion.

The objective of this study was to explore a new affective approach to projective mapping, with bread as case study, basing product categorization on consumers' choice or preference, and to compare it to the classic preference mapping approach. This approach might provide information that is more realistic for product developers and marketers during the product development process and market launch.

2. Materials and methods

2.1 Samples

Eight commercial wholegrain, pan-loaf breads were used in the study, bought in supermarkets in the region immediately south of Oslo (Norway). Products differed in terms of brands, prices, mix of grains used and percentage of wholegrain (Table 1).

2.2 Descriptive Analysis with a trained panel

A trained panel of nine assessors at Nofima Mat (Ås, Norway) performed a sensory descriptive analysis according to a quantitative descriptive analysis inspired by QDA® with modifications, as described by Lawless and Heymann (2010) as generic descriptive analysis. The assessors were tested, selected and trained according to ISO standards (ISO, 1993) and the sensory laboratory used followed the ISO standards (ISO, 1988). Nofima's panel is a highly trained and very stable panel; the assessors are solely hired as tasters, with a part-time job; some of them have more than 20 years' experience. The panel performance is assessed frequently, and checked for every project. The specific attribute list for the bread was developed in a one hour pre-trial session using two extreme bread samples. After a pre-trial session, the attributes and definitions were agreed upon by the assessors: they were all able to discriminate among samples, exhibited repeatability, and reached agreement with other members of the group. The assessors agreed upon 25 attributes describing the bread samples: odour intensity, hue, colour intensity, whiteness, pore size (crumb), amount of seeds/fibres (crust), roughness, elasticity, strength, crumbling, cohesiveness (using the finger), acidic taste, sweetness, saltiness, bitterness, yeast flavour, grain flavour, nut/seed flavour, roasted flavour, rancid flavour, hardness, juiciness, roughness/coarseness, chewiness and stickiness. All attributes were evaluated on unstructured line scales with labelled endpoints going from "no intensity" to "high intensity". In a pre-test session, the assessors were calibrated on samples that were considered the most different on the selected attributes typical for the breads to be tested. Samples were served in transparent Ziploc® bags labelled with three-digit numbers. Tap water was available for palate cleansing. Two replicates were performed for each bread sample. All samples and replicates were served in randomized order following a balanced block experimental design.

2.3 Consumer tests

Two sessions, one week apart, were held with the same group of participants and the same eight samples at Nofima Mat (Ås, Norway). In the first session, consumers performed two "classic" PM tests: blind PM (tasting blind samples) and informed PM (tasting together with the pack). In the second session, consumers first rated blind overall liking followed by a PM task based on choice or preference in informed conditions (tasting together with the pack). In both sessions, new samples with new codes were delivered for the two tests; consumers had a minimum of 15 minutes' break between tests.

2.3.1 Consumers' sample

The consumers included in the study (n=50) were recruited from Nofima's consumer database and were frequent consumers of wholemeal bread (more than twice per week). The participants were between 34 and 64 years old (43 years on average). Each session lasted around 30 min (Figure 1).

2.3.2 Session 1 – Classic PM, blind and informed

All participants were instructed in the use of the PM technique with a descriptive step. The basics of the technique were explained to the participants through an example employing geometric shapes with different colours and patterns, without any reference to bread. After the explanation of the technique, the participants received an A2 sheet of paper to allocate the samples. Samples were allocated according to the principle that samples with similar characteristics should be placed close to each other, while different samples should be placed further away. Next, they had to write down all the terms they could think of in connection with each sample, or group of samples, on the sheet, next to the position of the respective samples (technique also known as ultra-flash profiling).

Blind PM

The eight bread samples were presented simultaneously for direct comparison. Each sample was presented in a transparent Ziploc® bag coded with a three-digit number on a sticker. This type of presentation facilitated the location of the samples on the A2 sheet. The participants had to observe, smell and taste the breads, and then place the samples on the A2 sheet. Once they decided on the positioning, they were tasked with writing the codes on the sheet, and write the terms describing the perceived characteristics of the sample or group of samples close to the corresponding code.

Informed PM

The participants simultaneously received the eight bread samples in the same way as in the blind test, but this time each with an accompanying scan of the original front-of-pack (FOP), printed in colour. All scans of the FOP had the same dimensions. The participants performed the test in the same way as the blind test, but this time they had to consider both the information received and the sensory characteristics perceived. As before, they had to position the codes of the samples on the A2 sheet, and write down the descriptive terms.

2.3.3 Session 2 (one week apart) – Blind overall liking rating and informed PM based on choice or preference (PM-C)

Blind overall liking rating

Consumers rated their overall liking using 9-point box hedonic scales. Samples were assessed in blind conditions, in a rotated presentation order, balanced for order and carry-over effects (Wakeling & MacFie, 1995).

Informed PM based on choice or preference (PM-C)

Samples were presented in the same way as in the informed PM (bread samples with an accompanying front-of-pack), but with different codes. The instructions of this test differed from the "classic" PM approach in the way in which consumers had to base their categorization and sample allocation. Instructions were as follows (including underlining and capitals): "*Please evaluate the samples and look at the packs and position them on the sheet according to their differences and similarities* <u>*basing your criteria on what you would choose, thinking about different food occasions*</u>. *Place them on the sheet in such a way that two samples are close to each other if they're SIMILAR WITH REGARDS TO YOUR PREFERENCE and two samples are far from each other if they are DIFFERENT WITH REGARDS TO YOUR PREFERENCE*." As in the other two tests, after sample allocation, consumers had to write the codes of the samples on the A2 sheet together with descriptive terms.

These instructions were fine-tuned in a pilot test session before the main test (n=10). In the pilot, consumers went through the whole test (classic PMs, liking test, and PM-C). After the pilot trial, the researchers had an open discussion in which the consumers participated for feedback. For example, it was decided to add a phrase in the instructions stressing "*what you would choose, thinking about different food occasions"* to avoid consumers thinking they should just rank the samples from most to least preferred, basing their decision on only one consumption situation. In this way, they would understand that they could for example like two or more products equally, but could decide to consume them on different occasions or for different applications. In addition, pilot consumers suggested the categorization basis could be stressed by using capital letters: "*two samples are close to each other if they're similar with regards to your preference"* (and conversely). Based on the pilot it was also decided to include an example of a very different food category: sweet foods/desserts. They had different desserts, such as fresh fruit, yogurt, a gooey cake, etc. so they better understood the idea that it was possible to give multiple reasons for their choice.

2.3.4 Considerations on the experimental design

In session 1, the blind PM was done first and samples and map were taken away from the consumers when they had finished. The second part of the test was not explained to the consumers in advance; all they knew was that they were not done. After the 15-minute break, we instructed the consumers on how to do the informed PM test. The eight bread samples were different enough to be differentiated by means of direct comparison; however, they were eight (similar) slices of brown bread. It is very unlikely that the consumers remembered where they blindly positioned the eight samples from the blind PM to the informed PM, even if performed on the same day. The main driver for this experimental choice was that we wanted to keep the affective-based tests (Liking rating and PM-C) separated from the analytical approaches (classic PMs).

2.4. Data analysis

2.4.1 Preference mapping (sensory panel and consumer liking data)

An internal preference mapping was built through PLSR using the Consumercheck 1.4.2 open software tool. Consumer liking was used as the X matrix. The Y matrix were the sensory scores. Through this analysis, a score plot is obtained that visualizes how the products are related to each other in the space spanned by the first principal components, determined by consumer liking. The correlation loading plot shows how the variables of the X and Y matrices contribute to the common variation for each PC.

2.4.2 Analysis of the consumer test data

Analysis of variance (ANOVA) was performed on consumer overall liking scores considering consumer and sample as sources of variation. Mean ratings were calculated and significant differences were checked using Fisher's LSD test ($p < 0.05$). Agglomerative hierarchical clustering (HCA. Dissimilarity: Euclidean distance; Agglomeration method: Ward's method) was utilized as segmentation procedure in order to highlight groups of consumers with different liking patterns. Furthermore, an internal preference mapping was achieved via PCA (Principal Component Analysis) of a matrix of products x consumers to obtain a multidimensional representation of products and consumers in order to check against the clustering results (Varela, 2014). Analysis of variance (ANOVA) and Fisher’s test were also run for the clusters obtained, in the same way as above.

PM data in the three scenarios were collected as the X and Y coordinates of the samples on each consumer's individual map. A Multiple Factor Analysis (MFA) was performed considering the X and Y coordinates for the samples on each consumer's individual map as a group of variables (Pagès, 2005). Confidence ellipses were constructed as per Delholm et al. (2012). MFA was also carried out to compare the bread sample positions on the maps generated in the four evaluations. Values of RV coefficient were obtained for the purpose of comparing data from each session. RV ranges between 0 and 1; the closer to one, the greater the similarity between the configurations of the data tables.

To study if consumers grouped/mapped the samples differently in the three PM sessions, an MFA was conducted for the three tables for each consumer. Then the variability between the consensus of the three sessions was measured by the similarity index proposed in Tomic et al., 2015. In Tomic et al 2015, the SI was used to measure the variability to the consensus. Here we applied the same index for assessing the variability of each consumer across the different sessions. The similarity index (SI) for individual k in session i is computed as:

$$SI_{ki} = \frac{\|F_{ki} - F_k\|}{F_k}$$

Here $\| \|$ is the Frobenius norm, F_{ik} is the projected coordinates of consumer k from session i and F_k is the consensus of consumer k across the three sessions (i=1,2,3, k=1,2,…,n). The SI was computed for the consensus with A=2 components, hence there are two columns in F_{ki} and F_k. To measure how much the different consumers were influenced by the instructions, the average of SI over sessions was computed for each consumer. Higher SI values indicate that consumer maps were different in the different sessions, and that consumers were more affected by the instructions. There is no upper limit on SI, but a value > 1 indicates that residuals are larger than the variation between the samples within the consensus. The SI can also be computed for the complete data set in one session to measure the overall agreement of the consensus.

All the words provided by the participants in the descriptive step of the PM were analyzed qualitatively and differences were statistically checked, as follows: terms mentioned by at least 5% of the consumers were retained for further analysis (Symoneaux, Galmarini, & Mehinagic, 2012). The terms generated to describe the samples were grouped by consensus among two researchers, considering synonymous and derived words. The frequency table containing the terms was considered as a set of supplementary variables in the MFA of the PM data. The frequency of mentions was determined by counting the number of mentions of the same term in each session. Terms were grouped under three categories: *sensory, hedonics and usage & attitudes.*

Global Chi-square was used for testing the homogeneity of the contingency table of the terms generated in the descriptive step of the PM in the three scenarios (Symoneaux et al., 2012). When the initial Chi-square was significant, a chi-square per cell was done within each cell identifying the source of variation of the global Chi-square. This was run both for the individual terms and the three formed categories to compare the three scenarios.

The MFA analyses from the PM data were performed with the package FactoMineR (http://factominer.free.fr/) in R (version 3.2.2).

The chi-square per cell analysis was run with an XL macro as in Symoneaux et al. (2012).

The rest of the statistical analyses were run in XLSTAT, 2014, Addinsoft, New York

3. Results

It is important to point out that the objective of this methodological research was not to draw conclusions on the products themselves, but on how the different approaches to PM (analytical and affective) influenced the product descriptions and product choice information.

3.1. Overall Liking & liking patterns

Overall Liking (OL) significantly varied between bread samples (Table 2), ranging from 4.1 to 5.9. Preference responses are usually heterogeneous, and mean scores are not always representative of real preference patterns (MacFie, 2007; Felberg et al. 2010). Preference mapping approaches could be applied to understand consumer preference patterns, together with sensory data, to look for underlying dimensions that drive consumer preferences (Varela, 2014). In this first section, hierarchical cluster analysis (HCA) and the sensory description via generic descriptive analysis by the trained panel were combined to understand the liking patterns. Cluster analysis could be seen as "the lowest level of preference mapping" (Mac Fie, 2007).

HCA highlighted three clusters, one of them composed of only five consumers who rejected all samples (scores 4 and under). Assuming they disliked the general category under study, the analysis was continued on the other two clusters. Table 2 displays the distinct liking patterns of those two clusters. Although both groups of consumers rejected sample **B8**, liking patterns were clearly different. **B8** (barley, extra-coarse)**,** was described by the trained panel as having a rather strange, rancid flavor that may explain the general consumer rejection.

Cluster 1 discriminated less among samples. They rejected **B8** and did not present significant differences in overall liking among the rest of the samples; they were fairly open to any kind of bread but slightly preferred whiter, more cohesive breads.

Consumers in cluster 2 on the other hand, had more defined preferences, favouring dark, rough breads, and rejecting whiter, less coarse varieties. Samples **B1** (wholegrain, half-coarse) and **B5** were most liked and were described as having an intense odour, bitter, with nut/seed and roasted flavour, rough, with large pores, and dark. They were followed in liking rating by **B2** and **B7** (rye, extra-coarse), described as chewy, rough, sweet, roasted, dark and strong. Consumers in cluster 2 clearly rejected **B3** and **B4** (whiter, cohesive, sticky, crumbling, with yeast taste, grain taste and salty), added to the rejection of **B8**.

These liking patterns could be observed in the internal preference map (Figure 2).

In the following sections, the obtained two clusters will be explained by the descriptive data obtained by PM with consumers, to contrast with the interpretation provided by the trained descriptive panel. The conclusions that can be drawn with preference mapping approaches, combining classic descriptive data with overall liking, are limited to the sensory drivers of liking or disliking. The use of projective techniques such as PM permits understanding consumer perception beyond its sensory elements (e.g. attitudes, usage, affective terms), potentially revealing other reasons behind the affective response patterns (Ares et al., 2011; Varela & Ares, 2012).

3.2. Classic PM vs the new affective approach for understanding consumers' perception

3.2.1. Perceptual spaces – spatial configurations

Comparisons of the four evaluations

Sample configurations in the four tasting instances (descriptive analysis with the trained panel and the three PMs with consumers) were highly correlated, with RV coefficients ranging from 0.86 to 0.97. The generic descriptive analysis by the trained panel presented the lowest RVs with respect to all the PM scenarios, but still good enough (0.86). This can also be appreciated from the superimposed representation of the samples in the multiple factor analyses (Figure 3). For most of the samples, generic descriptive analysis was further away in the perceptual space to the consensus, but retained a similar relative position between samples. These results suggest that consumers may have a similar response regardless of whether they are assessing products blindly or informed, and even when basing the evaluation on their preference rather than on the products' descriptive characters. Moreover, the high correlations with the generic descriptive analysis indicate that the assessments are mostly based on sensory aspects.

In the descriptive step of blind PM, consumers generated a total of 75 different terms to describe the sample set, comprising mainly sensory terms (47) but also hedonic terms, and some related to usage and attitudes. In the descriptive step of the informed PM, consumers also generated 75 different terms in total, again including a majority of sensory terms (42) and some hedonic terms, as well as terms related to usage and attitudes. The fact that consumers focused more on sensory cues to describe similarities and differences among the samples rather than on usage or other elements accords with the high correlation obtained with the generic descriptive analysis and both classic PM tests.

In the descriptive step of the PM based on choice or preference, consumers generated approximately the same number of different terms in total (78); however, in this scenario the number of sensory terms was significantly lower (28), as highlighted by the chi square per cell analysis, and the description was more focused on the usage and attitudes category of terms (39). This shows that although the positioning of the products in the perceptual space might have been similar, consumers' associations when thinking about their preference or choice for different consumption occasions was different, and primarily driven by usage and the situation rather than by specific sensory cues. It should be noted that the PM-C instructions and dessert example primed consumers to think about usage and situations. Despite this, consumers could have used a similar number of sensory terms, which they did not. In a way, that was the idea behind the new approach: to prime them to be more specific about diverse drivers of their choices, going beyond the sensory experience, while also trying to retain the spontaneity of the projective technique as a basis.

Blind PM

Figure 4 shows the perceptual spaces as described by the two first dimensions of the MFA of the two classic PM in both scenarios (blind and informed). In the blind PM

(Figures 4 a1 and a2), the two first dimensions of the MFA display 50% of the variability of the original data. Considering together the samples' configuration (Figure 4 a1) and their description (Figure 4 a2), the breads were grouped mainly based on cereal type (oats, rye, barley, with wholegrain and combinations in the centre of the map), as well as fibre content and perception of healthiness. Consumers perceived the samples described as coarser and with a healthier taste (**B7**, **B5**, **B1**), while they associated more standard or ordinary traits with the softer samples on the other side of the first factor.

Informed PM

In the informed, classic PM: it is clearly visible from the sample configuration (Figure 4 b1) that the information polarized the results obtained for sample **B8**, which was separated from the rest of the samples in the consensus configuration. Evidently, the unique characteristics of this sample, particularly the "off-flavour" described by some consumers in the blind PM evaluation (Figure 4 a2) – in line with the "rancid" in the generic descriptive analysis – made more sense in consumer minds when knowing more about this bread. They mentioned the base cereal (barley and claims), focused more on describing the bad, off-taste, and mapped it further away from the rest. As B8 spans factor 2 of the MFA, the other samples do not show much variation in this direction. The first factor showed the variation of samples "from rye (**B7**) to oats (**B6**, **B4**)" with the wholegrain and mixes in the middle. However, variations in coarseness and darkness can be identified in this factor. The breads perceived as less coarse, or whiter are located towards the right of the plot. It is interesting to see that the information on the whole grain content did not noticeably affect the perception of coarseness, associated with **B7** and **B5** (extra coarse), but also with **B1** (half coarse).

PM based on choice or preference PM (PM-C)

Figure 5 displays the perceptual space obtained in the PM-C in informed conditions, as described by the two first dimensions of the MFA. Although the relative positioning of the samples in the spatial configuration was not essentially changed, enhanced discrimination between the products can clearly be observed in this scenario. Samples

B6 and **B4,** both made mainly with oats, were the only ones not discriminated in this tasting instance. In the PM-C, consumers used overall more words, and fewer words related to sensory descriptions. The extra information obtained with this type of PM approach can be appreciated in Figure 5 by interpreting the particular description of each sample (descriptive step), which can also be used to better understand the liking patterns as highlighted by consumers. For example, Cluster 2 preferred samples **B1**, **B2**, **B5** and **B7**, described in PM-C as dark, tasty, with good texture, a good/exciting taste, with corn, seeds and taste of seeds, sour, coarse, heavy, satiating, rich in fibre, healthy, sporty, for adults, of a well-known brand, rather expensive, good for dinner, with soup or cheese, and that they would buy them. On the other hand, consumers in Cluster 1 tended to like more chewy breads with a smooth surface, without whole seeds, less coarse, with oats, less tasty or even bland, good when toasted, a low price, everyday bread, for packed lunches, easily eaten, for families, for children. Meanwhile, these characteristics were rejected by cluster 2. The PM-C also helped to further understand the rejection of **B8** by all consumers. It was described as not attractive, with a bad, strange taste, off-flavour and odour, bitter, fluffy and porous and it was perceived as unhealthy; consumers stated they would not buy this kind of bread. This supports the idea of the different consumers' description in this case, driven by the usage occasions and the situation, and only a few important sensory cues.

Descriptive step

Table 3 shows the list of terms mentioned by consumers in the three PM scenarios together with the Chi Square per cell analysis. The terms included in the analysis were the ones cited by at least by 5% of the consumers of one product.

With respect to the ***sensory terms*** generated, even if there was a comparable number of different terms cited in the blind (47) and informed PM (42), the frequencies of citation were in general higher in the blind tasting, as consumers relied mostly on the sensory characters when explaining their maps. The terms mentioned most frequently in the blind PM (with more than 40 mentions) were: bland, bright colouring, coarse, corn, dry, seeds/taste of seeds. In the informed PM, the sensory terms were fewer in total, but the most frequently mentioned were largely the same; however, juicy and smooth surface also became important terms used to describe the samples in this scenario. In the PM-C, the total number of sensory terms was significantly lower (28), as highlighted by the chi-square per cell analysis, and the terms elicited by consumers with high frequency were fewer. However, the words bland, corn and dry continued to be mentioned more than 40 times, but significantly less frequently than in the blind scenario. However, coarseness was mentioned significantly more frequently, going from 44 mentions in the blind PM to 106 mentions in the affective approach (PM-C); this suggests that coarseness may have been one of the most important drivers of product differentiation when considering choices in this particular sample set.

The ***hedonic terms*** category was the one with fewest distinct terms generated by consumers in the three PMs, and the frequencies were also lower. In general, in the blind PM there were significantly more terms that expressed liking or disliking of some sensory characteristics, such as: exciting appearance, good smell, standard appearance and standard texture; however, the number of mentions was low (25 or less). The hedonic term most mentioned in the three PM was good/exciting taste, but there were no differences between them (86-101 mentions). It is quite interesting how two of the hedonic terms significantly increased in the PM-C. Bad taste and would not buy/eat/uninterested became very important in the affective approach, which suggests that consumers were more prone to express their opinions with regards to disliking when grouping the samples based on what they would actively choose (in a real-life scenario).

The category of descriptions on ***usage & attitudes*** was more heavily influenced by the scenario. The number of different terms generated in total more than doubled in the affective approach to PM (from 15 in blind to 39 in the affective approach), and the frequencies of mention of usage & attitudes terms were significantly higher. The terms generated included: target consumers (for kids, for adults, for family), consumption occasions (for breakfast, lunch, dinner, everyday bread, for packed lunches, for sport), food pairings (for soup, with cheese, with toppings, with jam, versatile), health-related properties (healthy, satiating, weight-reducing), references to the brand (good label, standard label), and to the price (expensive, low price). It is interesting to highlight how the price references were almost non-existent in the classic PM scenarios (both blind and informed), and how the references to healthiness increased significantly, apart from focusing much more on the possibilities of product usage .

Chi square per cell was also run on the term by product matrix in each scenario, to being able to highlight the different profiles of each sample (data not shown). As stated above, the main objective of this paper was not to describe the samples; nevertheless the study shows that the terms generated by each individual product in the affective PM highlight the important attributes for each sample in the light of the different preference patterns. For example, **B8** was associated significantly more frequently with the terms would not buy, bad taste, weird taste, off-flavour, sour taste and non-informative label. Hence it this explains why the product was rejected by most consumers, highlighting the drivers of disliking. On the contrary, **B5**, the bread liked by both groups of consumers, was associated more frequently with terms such as with a good/exciting taste, tasty, with good smell and good-tasting crust, and consumers found it good both as bread for packed lunches and sporty. In terms of coarseness, it was significantly associated with this concept, but not significantly different to **B7**, which was viewed to a significantly greater degree as a dark bread, for adults and highly satiating. This suggests that **B5** could be a good option for both clusters within the coarser breads, while **B7** was very well-liked by Cluster 2 but within the less liked samples in Cluster 1.

3.4. Consumers' individual behaviour in the different PM scenarios

A natural question that might be raised at this point is how different consumers, or groups of consumers, reacted to the change in PM scenario. When comparing how samples were located in the perceptual spaces by both liking clusters in the different tests, they were also very similar; for example, comparing the relation of the perceptual spaces obtained by clusters 1 and 2 in the PM-C, RV was 0.882. Something similar happened when comparing the outcomes for the same cluster throughout scenarios; for instance, Cluster 1 had an RV of 0.828 between PM blind vs. PM-C. These results showed that the maps obtained for the groups with similar liking patterns were quite stable throughout different PM tests. However, that was not necessarily the case when studying consumers' individual behaviour. Some of the consumers changed their maps drastically from one scenario to another, while a few others maintained very stable mapping structures throughout assessments. Figure 6 presents the MFA plots comparing the three evaluations for the two consumers that presented the best (C118) and worst (C121) agreements between sessions. Consumer C118 performed a highly similar comparative allocation of the samples in the three perceptual spaces, with high RV coefficients (RV inf-blind= 0.71; RV choice-blind= 0.76; RV inf-choice= 0.86). On the contrary, the perception of the samples for consumer C121 shifted notably from scenario to scenario, with very low RV coefficients (RV inf-blind= 0.1; RV choice-blind= 0.1; RV inf-choice= 0.04). To obtain an overall view of the consumer sample, the SI (similarity index) coefficients were calculated for each of the participants (Tomic, Berget & Naes, 2015). SI takes a value of zero when configurations are the same as the consensus scores; the higher the value, the lower the similarity. Figure 7 shows the distribution of SI values for all the consumers, ranging from 0.47 to 1.11. Most consumers had SI values between 0.6 and 0.8. Few consumers have a much worse or much better fit than the rest, suggesting that there were relatively small individual differences.

4. General Discussion

The fact that consumers might react similarly when mapping products based on their preferences or choice as compared to when mapping products based on the products' descriptive similarities or differences, and that these mappings might be mostly based on the sensory aspects, was initially unexpected. Carrillo et al. (2012a, 2012b) had similar findings when comparing results of classic blind and informed PM on biscuit samples, hypothesizing that product information is in fact a ‘‘modulator’’ of consumer perception, meaning that the perception is basically one which would be modulated depending on the context experienced by the consumer. In this way, individual sample characterization would vary within the perceptual space but the sample multivariate structure (distance and relative positioning among products) would not vary dramatically. The same authors found that the observed changes presented a sample-dependent effect. This was also the case in the present work. When looking at figures 4 and 5, it is evident that samples **B2**, **B5** and **B8** shifted positions considerably more than the other samples, while the overall structure of sample configuration remained stable. In particular, **B8** was assessed as very different from the rest (polarizing effect) when assessed with information, both in the informed PM and in the PM-C. This shift may have occurred because it was the only sample that contained barley and because of its on-pack nutritional and health claims (B-glucans, lower cholesterol, long-lasting satiety). Carrillo et al. (2012a) mentioned a sample-dependent change in perception linked to nutritional and health claims, particularly when those claims were not completely understood by consumers. Added to this, other authors have highlighted the importance of the fit carrier-claim (Krutulyte et al., 2011), and how the perceived carrier-ingredient fit is related to the familiarity with the combination and to the healthiness of the carrier food (Carrillo et al., 2012b). Barley, albeit not an unknown bread ingredient for Norwegian consumers, has been re-introduced in the Norwegian market in many new products accompanied by the communication of various health and nutritional effects. B-glucan is also quite a new functional ingredient for the Norwegian market.

The reported stability of sample configurations in blind and informed conditions, also demonstrated by the present study, and the modulator effect of the context of the test, make sense in an analytic descriptive framework. This is because consumers use the available information to sort samples in a bi-dimensional perceptual space which would subsequently be modified by the extra information received through the pack. Further, the results of this and previous works using PM in different scenarios suggest that this basic perceptual structure in consumers' minds would be determined primarily by the product sensory cues and modulated by the extrinsic product information. This modulation is expressed by tweaking the map, and mainly by using specific and distinct characteristics in the descriptive step. It would be worthwhile to study the effect (or absence of an effect) of this modulation in other type of studies, for example in conjoint approaches, as compared to PM, looking into the interaction of intrinsic and extrinsic product cues. In those tests, the information is usually displayed on a computer screen, showing all variables with the same salience, something that could potentially lead to an overestimation of the influence of certain parameters on food choice, as previously suggested by Varela et al. (2014).

The idea behind the method suggested in this paper and some of the results of the present study were presented in Eurosense 2014 and not published until now for a range of reasons. In the meantime, we had the chance to conduct a second study using PM-C and to compare it to CATA, to evaluate consumers' perception of a complex set of stimuli such as aromatically enriched wines. In that recently published work (Lezaeta et al., 2017), working with 150 consumers, we observed that both consumer-based methods highlighted the positive effect of aromatic enrichment on consumer perception and acceptance. However, PM-C generated a very detailed description in which consumers focused less on the sensory aspects and more on the usage, attitudes, and reasons behind their choices, providing a deeper understanding of the drivers of liking/disliking of enriched Sauvignon Blanc wines. This new work confirmed what we suggested in the proof of principle, which we now elaborate on in this work.

However, prior to these two studies, there was no experience with changing the cognitive framework of Projective Mapping from an analytic mapping to an affective mapping, and our results suggest that consumers would be performing a sort of "preference mapping in their heads". To accomplish this aim, they would first map the products, as they would do in a classic PM, and they would subsequently state their preferences via the descriptive step, for example by describing usage and attitudes characteristics in considerable detail. More work would be needed on this technique to assess if this can be generalized to other cases. It is also possible that the affective frame of mind allowed for better differentiation between the samples, through a combined effect of the modulation of the extrinsic characteristics and the personal meaning added to the different product dimensions (hedonic perception, usage, attitude, brand perception, etc.). Indeed, in Lezaeta et al. (2017), we saw that – compared with CATA – PM-C stretched the perceptual space further, with PM-C discriminating better among the wine samples.

In the 1998 paper by King et al., comparing free and structured projective mapping (with liking as one of the axes) for identification of similarity-of-use of snack bars, they did not obtain a better sample discrimination through the structured PM. It is possible that a too-structured mapping scenario, with predefined categories, prevented consumers from freely expressing their perceptions, sorting the products into relatively obvious groups rather than detailing their hedonic perception. Torri et al. (2013) studied how different groups of consumers realised a classic PM test with wines, where consumers' product differentiation was poor. They separated the consumers into three groups depending on their performance and concluded that increased differentiation ability was observed among those consumers able to match the duplicate samples in the PM test, and that their main mapping dimension was highly correlated to their liking. Even if consumers were asked to describe the samples and no indication of using liking as criteria was given, it is possible that the high complexity of the samples pushed some consumers into using their hedonic perception as a basis for categorization. Those consumers were able to improve discrimination, which would be in agreement with what was reflected in our work.

The descriptive step in the affective approach to PM provided a much richer description than the classic approach in terms of preference drivers. Consumers expanded on the reasons behind sample categorization and their choices, covering things such as target consumers, consumption occasions, possibilities of usage, food pairings, health-related properties, brand associations and references to the price and willingness to buy/not buy.

In this scenario, consumers also highlighted their rejection or disliking drivers in greater depth.

5. Conclusions

The results of the perceptual spaces obtained in this work comparing PM in blind and informed conditions were quite comparable, suggesting that sensory cues were the main driver in the categorization. In the choice-based PM, consumers focused less on the sensory aspects and more on usage & attitudes, generating more detailed descriptions. In this way, the affective approach to PM provided an enhanced understanding in terms of the drivers of liking/disliking, making it a promising potential tool for category and market exploration.

The limited number of consumers used in this study (n=50) did not permit drawing any conclusions on implications for the bread category in the Norwegian market. This was not an objective of this work, but rather a proof of principle of the approach. The clear differences found when comparing PM scenarios make the data strong enough from a methodological perspective, suggesting that this new approach to PM could add interesting information on consumers' drivers for liking and reasons behind their choices. More research is needed on further product categories to further improve understanding of the complete picture.

It is in fact interesting how PM-C allowed for this "unfolding" in a seemingly two-step processing and conveying of the information: first, a sensory description, followed by an in-depth hedonic and behavioural description. This phenomenon deserves further research.

As pointed out by some recent methodological studies in classic PM (Varela et al., 2014; Vidal et al., 2016; Varela et al., 2017) it would be also worth following up the individual differences and group behaviour in the PM-C.

Acknowledgements

We would like to thank Merete Rorvik and Heidi Birkelund from Coop Norge for support with sample selection. The authors would also like to express their gratitude for the financial support received from the Norwegian Foundation for Research Levy on Agricultural Products FFL through the research program "FoodSMaCK, Spectroscopy, Modelling and Consumer Knowledge" (2017-2020), and the Research Council of Norway through the RapidCheck project. We also wish to thank the European Commission for its support through the Marie Curie Actions Intra European Fellowship (IEF), call FP/-PEOPLE-I2012-IEF – project title ''Innovative Methodologies for New Food Product Development: combining Sensory Science and Experimental Economics – NEFOMET".

References

Ares, G., Deliza, R., Barreiro, C., Giménez, A. & Gámbaro, A. (2010). Comparison of two sensory profiling techniques based on consumer perception. *Food Quality and Preference,* 21, 417–426.

Ares, G., Varela, P., Rado, G. & Giménez, A. (2011). Identifying ideal products using three different consumer profiling methodologies. Comparison with external preference mapping. *Food Quality and Preference,* 22, 581–591.

Ares, G. & Varela, P. 2014. Comparison of Novel Methodologies for Sensory Characterization. In: Novel Techniques in Sensory Characterization and Consumer Profiling. Eds: P. Varela and G. Ares. pp. 365-389. CRC Press: Boca Raton, FL.

Asioli D, Næs T, Granli BS, et al. (2014) Consumer preferences for iced coffee determined by conjoint analysis: an exploratory study with Norwegian consumers. International Journal of Food Science & Technology, 49(6), 1565–1571, Available from: http://dx.doi.org/10.1111/ijfs.12485.

Asioli D, Varela P, Hersleth M, et al. (2017) A discussion of recent methodologies for combining sensory and extrinsic product properties in consumer studies. Food Quality and Preference, 56, 266-273

Jaeger S (2006) Non-sensory factors in sensory science research. Food Quality and Preference, 17(1–2), 132–144, Available from: http://www.sciencedirect.com/science/article/pii/S0950329305000285.

Maslow AH (1954) Motivation and personality. New York: Harper & Row.

Carrillo, E., Varela, P. & Fiszman, S. (2012a). Packaging information as a modulator of consumers' perception of enriched and reduced-calorie biscuits in tasting and non-tasting tests. *Food Quality and Preference,* 25, 105–115.

Carrillo, E., Varela, P. & Fiszman, S. (2012b). Effects of food package information and sensory characteristics on the perception of healthiness and the acceptability of enriched biscuits. *Food Research International,* 48, 209–216.

Dehlholm, C., Brockhoff, P. B. & Bredie, W. L. P. (2012). Confidence ellipses: A variation based on parametric bootstrapping applicable on Multiple Factor Analysis results for rapid graphical evaluation. *Food Quality and Preference,* 26, 278-280.

Dehlholm, C. (2014) Projective Mapping and Napping. In: Novel Techniques in Sensory Characterization and Consumer Profiling. Eds: P. Varela and G. Ares. pp. 365-389. CRC Press: Boca Raton, FL.

Felberg, I., Deliza, R., Farah, A., Calado, E. & Marino-Donangelo, C. (2010) Formulation of a soy-coffee beverage by response surface methodology and internal preference mapping. Journal of Sensory Studies 25, 226–242.

Guerrero, L., Claret, A., Verbeke, W., Enderli, G., Zakowska-Biemans, S., Vanhonacker, F., Issanchou, S., Sajdakowska, M., Granli, B.S., Scalvedi, L., Contel, M., Hersleth, M. (2010). Perception of traditional food products in six European regions using free word association. *Food Quality and Preference,* 21, 225–233.

ISO (1988). International Standard 8589. Sensory Analysis – General Guidance for the Design of Test Rooms. [Ref. No. ISO 8589:1988 (E)] International Organisation for Standardisation, Genève (1988)

ISO (1993). International Standard 8586–1. Sensory analysis – Methodology – general guidance for the selection, training and monitoring of assessors. [Ref. No. ISO 8586–1:1993 (E)] International Organisation for Standardisation, Genève.

King, M.C., Cliff, M.A. & Hall J.W. (1998). Comparison of projective mapping and sorting data collection and multivariate methodologies for identification of similarity-of-use of snack bars. *Journal of Sensory Studies,* 13, 347–358.

Krutulyte, R., Grunert, K., Scholderer, J., Lähteenmäki, L., Hagemann, K., Elgaard, P., et al. (2011). Perceived fit of different combinations of carriers and functional ingredients and its effect on purchase intention. *Food Quality and Preference,* 22, 11–16.

Lawless, H. T. & Heymann, H. (2010) Sensory Evaluation of Food: Principles and Practices. Second Ed. Springer Science & Business Media. 596p. ISBN 978-1-4419-6487-8

Lezaeta, A., Bordeu, E., Næs, T., & Varela, P. (2017). Exploration of consumer perception of Sauvignon Blanc wines with enhanced aroma properties using two different descriptive methods. *Food Research International*. In press.

MacFie, H. 2007. Preference mapping and food product development. In: Consumer-Led Food Product Development. (ed. H. MacFie), pp. 407-433. Woodhead Publishing Ltd. Cambridge, UK.

Miraballes, M., Fiszman, S., Gámbaro, A. & Varela, P. (2014). Consumer perceptions of satiating and meal replacement bars, built up from cues in packaging information, health claims and nutritional claims. *Food Research International,* 64, 456–464.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire valley. *Food Quality and Preference,* 16, 642–649.

Symoneaux, R., Galmarini, M.V. & Mehinagic, E. (2012) Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. *Food Quality and Preference,* 24, 59–66

Tomic, O., Berget, I., Næs, T. (2015) A comparison of generalised procrustes analysis and multiple factor analysis for projective mapping data. *Food Quality and Preference,* 43, 34–46

Torri, L., Dinnella, C., Recchia, A., Naes, T., Tuorila, H., Monteleone, E. (2013) Projective Mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference,* 29, 6–15.

Valentin, D., Chollet, S., Lelièvre, M., and Abdi, H. (2012). Quick and dirty but still pretty good: A review of new descriptive methods in food science. *International Journal of Food Science and Technology,* 47, 1563–1578.

van Kleef, E., van Trijp, H. C. M., & Luning, P. (2006). Internal versus external preference analysis: An exploratory study on end-user evaluation. *Food Quality and Preference*, 17, 387–399.

Varela, P. (2014). Application of multivariate statistical methods during new product development – Case study: Application of principal component analysis and hierarchical cluster analysis on consumer liking data of orange juices. In: Mathematical and Statistical Methods in Food Science and Technology. Eds: Granato, D. & Ares, G. pp 187-200. John Wiley & Sons Ltd. DOI: 10.1002/9781118434635

Varela, P. & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International,* 48, 893–908.

Varela, P., Antúnez, L., Silva Cadena, R., Giménez, A. & Ares, G. (2014). Attentional capture and importance of package attributes for consumers' perceived similarities and differences among products: A case study with breakfast cereal packages. *Food Research International,* 64, 701–710.

Varela, P., Antúnez, L., Berget, I., Oliveira, D., Christensen, K., Vidal, L., ... & Ares, G. (2017). Influence of consumers' cognitive style on results from projective mapping. Food Research International.

Wakeling, I.N. & MacFie, H.J.H. (1995) Designing consumer trials balanced for first and higher orders of carry-over effect when only a subset of k samples from t may be tested. *Food Quality and Preference,* 6, 299–308.

Vidal, L., Antúnez, L., Giménez, A., Varela, P., Deliza, R., Ares, G. (2016) Can consumer segmentation in projective mapping contribute to a better understanding of consumer perception? *Food Quality and Preference,* 47, 64–72.

Table Captions

Table 1.- Bread samples included in the research

Table 2.- Mean OL ratings and Fisher LSD (n=50, Analysis of the differences between the categories with a confidence inteRVal of 95%)

Table 3.- Descriptive step in the three PM assessments. Chi square per cell analysis. The analysis was run in the complete data table. Data are displayed in three groups (sensory terms, hedonic terms and usage and attitudes terms) for better understanding.

(+) or (-) indicate that the observed value is higher or lower than the expected theoretical value. *** p < 0.001,** p < 0.01 and * p < 0.05; effect of the chi square per cell

Figure captions

Figure 1.- Workflow of experiments

Figure 2.- Internal preference map, (a) product plot and (b) consumers and attributes plot

Figure 3.- Superimposed MFA representation of the eight samples. Each sample is represented by four points, corresponding to the four assessment instances: QDA (generic descriptive analysis), PM Blind, PM Informed, PM Choice). The consensus representation is represented for each of the samples as the central point.

Figure 4.- Multiple factor analysis of the data obtained in the two classic PM scenarios. (a1) Representation of the samples in the PM Blind; (a2) Representation of the terms in the PM Blind; (b1) Representation of the samples in the PM Informed; (b2) Representation of the terms in the PM Informed.

Figure 5.- Multiple factor analysis of the data obtained in PM based on choice. Representation of the samples (left) and the terms (right)

Figure 6.- Superimposed MFA representation of the eight samples, corresponding to the three PM evaluation instances, for two individual consumers. Consumer with best agreement on the left (RV inf-blind= 0.71; RV choice-blind= 0.76; RV inf-choice= 0.86) and the consumer with the worst agreement on the right (RV inf-blind= 0.1; RV choice-blind= 0.1; RV inf-choice= 0.04).

Figure 7.- Barplot showing the similarity index (SI) for all consumers. The values are sorted so that the consumers on the very left have the smallest variation across the different sessions, whereas the consumers on the very right have large variation across the sessions.