Food Quality and Preference 61 (2017) 6–14

Contents lists available at ScienceDirect

Food Quality and Preference

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

Pivot profile method: What is the influence of the pivot and product space?

Maud Lelièvre-Desmas[a,*], Dominique Valentin[b], Sylvie Chollet[a]

[a] Institut Charles VIOLLETTE (ICV) EA 7394, ISA, Univ. Lille 1, INRA, Univ. Artois, Univ. Littoral Côte d'Opale, F-59000 Lille, France
[b] Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRA, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France

A R T I C L E I N F O

Keywords:
Pivot profile
Reference
Descriptive method
Free description, beer

A B S T R A C T

The Pivot Profile© (PP) was proposed recently by Thuillier, Valentin, Marchal, and Dacremont (2015) as an alternative to classical descriptive methods. Its principle is to describe each product by comparing it to a stable reference (called Pivot). While the method seems promising there is little data available and some issues still need to be examined. This paper proposes to evaluate two of these issues: the effects of the similarity within the product space and of the choice of pivot. We compared the pivot profiles obtained for three different sets of beers, varying in their within-set sensory similarity, using different pivots. We found that PP results are more influenced by the within-set similarity than by the choice of the pivot. We suggest that the PP method is more suitable for restricted product spaces in terms of sensory characteristics, and that the creation of a "central product" as the pivot can be a good option when the type of products allows it. However, further studies need to be conducted to assess PP in terms of assessors' repeatability and consensus, as well as to propose alternative statistical analyses that would take into account the individual PP data.

1. Introduction

In response to industrial demand to develop fast and cost effective methods to describe product sensory attributes, new sensory tools have been described in the literature under the name of rapid methods (Valentin, Chollet, Lelièvre, & Abdi, 2012; Varela & Ares, 2012). One of the most recent tool was proposed under the name of Pivot Profile© (PP) by Thuillier et al. (2015) for the description of champagne wines. The idea behind PP is to use a comparative strategy like in Flash profile (Dairou & Sieffermann, 2002; Delarue, 2014), sorting task (Chollet, Valentin, & Abdi, 2014) or projective mapping (Dehlholm, 2014; Risvik, McEwan, & Redbottena, 1997) but comparing each product to a stable reference instead of comparing all products together. This characteristic is common with Polarized Sensory Positioning (PSP, Teillet, Schlich, Urbano, Cordelle, Guichard, 2010). In PSP, assessors are asked to evaluate the distance between a sample and three references or poles whereas in PP they have to describe the differences between the sample and the reference (pivot). One of the main interests of these two reference-based methods is that they allow data aggregation as the samples do not need to be presented all at the same time as for the other comparative methods previously quoted (Valentin et al., 2012). The main difference between the two methods is that PSP does not rely on language and so it provides only a positioning of samples, when PP provides both a positioning and a description of the samples.

Practically in PP assessors are provided with pairs of products including the reference or pivot (clearly identified as such) and the product to be evaluated. Assessors are asked to observe, smell and taste the pivot and the product and to write down each attribute that the product has in smaller or larger amount than the pivot product (*e.g.* less sweet, more astringent). Data analysis begins by regrouping synonyms and optionally regrouping the terms by categories. Then, negative and positive frequencies are computed for each term and each product, and the negative frequency is subtracted from the positive frequency. The resulting score is finally translated so as to obtain positive scores only. The final matrix is submitted to correspondence analysis (CA) to obtain a sensory map of the products.

PP seems to be very promising as it allows for a fast description of products with the possibility of aggregating data across sessions. However, so far very little data are available with only two published articles at that day. Thuillier et al. (2015) reported that champagne descriptions obtained *via* PP were coherent with what was *a priori* known about these wines and Fonseca et al. (2016) demonstrated with chocolate ice-creams that PP has high analytical and discriminative power compared to comment analysis. So, additional work is clearly needed to confirm the potential of this method. A first issue with PP is the choice of the pivot product. In a series of simulations Thuillier et al. (2015) demonstrated that "*the choice of the pivot does not impact product description space in a dramatic way*" (p. 72). Yet this still needs to be verified with real products and assessors. The difficulty is that if the pivot is too neutral compared to the products to be described, assessors

* Corresponding author at: ISA Lille, 48 Boulevard Vauban, 59046 Lille Cedex, France.
E-mail address: maud.desmas@yncrea.fr (M. Lelièvre-Desmas).

http://dx.doi.org/10.1016/j.foodqual.2017.05.002
Received 22 November 2016; Received in revised form 3 May 2017; Accepted 3 May 2017
Available online 05 May 2017
0950-3293/

might have difficulties finding terms that are less than the pivot and inversely if the pivot is too caricatural assessors might have difficulties finding terms that are more than the pivot. The ideal solution thus is to have a pivot with a central position compared to the product space to be described. Thuillier et al. (2015) suggest creating a "central product" by blending all products to be described but specify that this option is possible "*only for liquid, semi-liquid or powder products that can be easily mixed*" (p. 72). However, the author did not test this idea. The same issue holds for PSP as a minimum of two poles need to be chosen. Teillet, the author of the method himself, indicates that "the choice of the poles seems to be a critical point of PSP" (Teillet, 2014, p. 265). Recent publications on this issue reports that small changes in the choice of the poles do not lead to relevant changes in product configurations but it seems crucial that the poles should reflect the main sensory characteristics responsible for the expected similarities and differences among the products to be evaluated (Ares et al., 2015; De Salamando, Antúnez, Torres-Moreno, Giménez, & Ares, 2015). Another issue is related to the similarity within the product space (within-set similarity). Most rapid methods do a good job at describing and discriminating among products as long as the differences between products are not too small. When the differences are too subtle, classical descriptive analysis is generally more efficient (Antúnez, Vidal, de Salamando, Giménez, & Ares, 2017).

The aim of this paper was to evaluate the potential of PP to describe complex products and to evaluate the effect of within-set similarity and of choice of pivot on PP performance. We selected three sets of beers as an illustration of complex products (Fig. 1). The first set is composed of very different beers (low within-set similarity), the second set has an intermediate within-set similarity and the third set is the less similar one. Each set of beers was evaluated in comparison to one, two or three different pivot products.

2. Material and methods

For assessing the effect of the choice of pivot we evaluated Sets 2 and 3 respectively with several pivots (blue lines in Fig. 1). To obtain a central product we followed Thuillier et al.'s suggestion and created a blend of all beers of each set. For Set 2, we compared this central blend pivot (P1) with two extreme pivots which are real beers. For Set 3, we compared the central blend pivot (P1) with a central pivot beer (P2) which is a real beer quite similar to the beers of Set 3. If PP is sensitive to the choice of pivot, we expected to obtain different product descriptions of the beers depending on the pivot product inside each set.

For assessing the effect of within-set similarity on PP performance, we compared the results of the three sets when evaluated in comparison with their respective central blend pivot (P1, orange line in Fig. 1). If PP is sensitive to within-set similarity, we expected PP to give better performance with Set 1 (the less similar) than with Set 3 (the more similar).

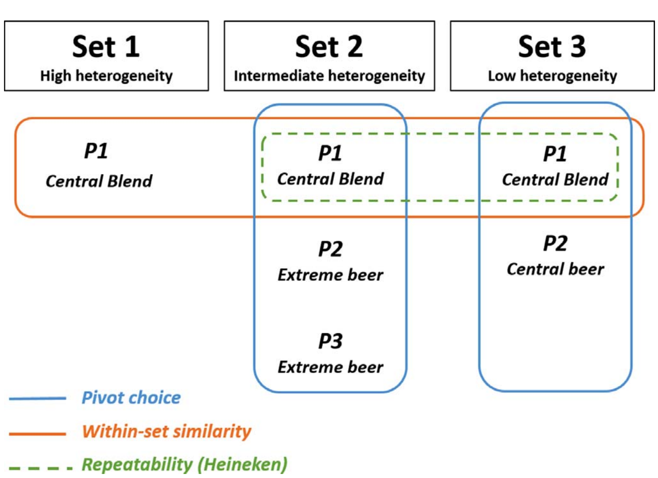

Fig. 1. Global schema of the experiment.

2.1. Assessors

In their simulations, Thuillier et al. (2015) evaluated the effect of panel heterogeneity as a factor that could impact PP outcome. They showed that several sets of heterogeneous individual descriptions lead to similar product descriptive spaces which suggests that this factor might not be a problem. However, they also mention that the heterogeneity they simulated was closer to the heterogeneity that could be expected from a trained panel and that a bigger effect could be expected with consumers. Based on this statement we used trained panelists as assessors to avoid additional noise in the data. So 11 trained assessors (6 women, 6 men aged from 28 to 59 years old) enrolled in a training program designed to produce beer trained panelists. They had been formally trained to evaluate different kinds of beers (including the beers studied here) one hour per week for an average of four years. The training consisted in detecting and identifying flavors in beer and evaluating the intensity of general compounds on a non-structured linear scale. They did not have any knowledge about PP and had never participated in a PP test before.

2.2. Beers

The choice of the beers was based on previous published and unpublished studies involving sensory profiling of a large number of beers, and on the extensive knowledge of the authors of the beer sensory characteristics grounded on more than 15 years of research activity on this topic. The beers of each set are detailed in Table 1. Set 1 (low within-set similarity) includes both blond and amber beers with various alcohol contents. Set 2 (intermediate within-set similarity) includes five blond beers: two of them are more aromatic and alcoholic (Leffe and Grimbergen) than the three others (Pelforth, Stella Artois and Heineken). Set 3 (high within-set similarity) is composed of five similar blond beers with very close alcohol contents.

For each set, the central blend pivot is made of an equi-volume

Table 1
The three sets of beers with their respective pivots.

	Beers	Color	Alcohol content (% vol.)
Set 1	Chti	Blond	6.4
	Chti	Amber	5.9
	Leffe	Blond	6.6
	Leffe	Amber	8.2
	Pelforth	Blond	5.8
	Pelforth	Amber	6.0
Pivot 1	*Blend of the 6 beers*	/	/
Set 2	Grimbergen	Blond	6.7
	Leffe	Blond	6.6
	Pelforth	Blond	5.8
	Stella Artois	Blond	5.2
	Heineken*	Blond	5.0
Pivot 1	*Blend of the 5 beers*	/	/
Pivot 2	*Affligem*	*Blond*	*6.7*
Pivot 3	*St Omer*	*Blond*	*5.0*
Set 3	1664	Blond	5.5
	Carlsberg	Blond	5.0
	Heineken*	Blond	5.0
	St Omer	Blond	5.0
	Stella Artois	Blond	5.2
Pivot 1	*Blend of the 5 beers*	/	/
Pivot 2	*Jupiler*	*Blond*	*5.2*

* The Heineken beer was presented two times to the assessors within Sets 2 and 3 (compared to Pivot 1 each time) to assess the repeatability of the method.

blend of the beers composing the set. Additional real beers were selected as other pivots for Sets 2 and 3. For Set 2, we chose two extreme pivots: P2 (Affligem) which is more similar to two beers (Leffe and Grimbergen) and P3 (St Omer) which is more similar to three beers (Pelforth, Stella Artois and Heineken). For Set 3, P2 (Jupiler) was selected as a central pivot beer which has similar sensory characteristics to the beers of the set. Thereafter in the text, the association of one set with one pivot will be named *trial*. All beers were presented in three-digit coded black glasses and served at 10 °C.

2.3. Procedure

Assessors took part individually in the experiment which was conducted in separate booths lighted with a neon lighting of 18 W with a red filter to mask the color differences between beers. Mineral water and bread were available for assessors to rinse between samples. Assessors could spit out beers if they wanted. The experiment was run in seven sessions, at the rate of one session per week. The evaluation order of each set was balanced among the assessors and for Sets 2 and 3 the order of the pivots was also balanced among the assessors. At the end of the last session, the Heineken was presented twice again to the assessors: one time in comparison to Pivot 1 of Set 2 and the other time in comparison to Pivot 1 of Set 3 (in a balanced order), in order to evaluate the repeatability of the method (dotted green line in Fig. 1).

During each session, assessors were successively presented with four or five pairs of beers composed of the pivot beer and the beer to be evaluated. A new glass of pivot beer was presented for each evaluated beer, to avoid any temperature or sparkling variation during the session. For each pair, assessors were asked to smell and taste both beers and to write down all terms that they perceived in lower or higher intensity in the evaluated beer than in the pivot beer. They were asked to use only descriptive terms without any sentence. Negative forms were not allowed since they could be turned into positive form (*e.g.* non-sparkling into flat).

2.4. Statistical analysis

The effect of within-set beer similarity and of choice of pivot were analyzed separately. For both effects we first compiled the terms generated for each set of beers. This step was operated manually by two of the three authors. Then the different grammatical forms of a same word were standardized (*e.g.* hoppy and hop).

2.4.1. Performance criteria

Then we evaluated the richness of vocabulary, the agreement between assessors and their repeatability by computing three criteria:

Number of terms. The average number of terms for each trial was computed across assessors and beers. To assess the effect of within-set similarity, we performed a two-way ANOVA with assessors as random factor and set as fixed factor (Set 1, 2 & 3). To assess the effect of choice of pivot, we performed two two-ways ANOVAs, one on Set 2 and one on Set 3. In both ANOVAs assessors were considered as random factor and pivot as fixed factor (P1, P2 & P3 for Set 2 and P1 & P2 for Set 3). When a significant effect of set or pivot was found, a Newman Keuls pair comparison test was performed with an α-level sets at 5%.

Global consensus index. First, for each beer and for each pair of assessors, we assigned a value of 1 to a term if it was used similarly by both assessors (positively or negatively), a value of 0 if the term was used by only one of the two assessors and a value of -1 if both assessors used it in different directions (one negative and the other positive). These values were then averaged and compiled in an assessor*assessor matrix for each beer. Each cell in this matrix represents the agreement between a pair of assessors for a given beer. Individual consensus scores were then obtained for each beer by summing the rows of the matrix and dividing the result by the number of assessor minus 1. The global consensus index was then obtained by averaging the individual consensus score across assessors and beers. This index varies from -1 to $+1$. The same ANOVAs as for the number of terms were carried out to evaluate the effect of within-set similarity and choice of pivot.

Repeatability index. First the similarity between the two replications of the Heineken beer within Set 2 and Set 3 is computed for each assessor using the same algorithm as for the consensus index. Then the repeatability index is obtained as the average of the similarity across assessors for each set. To assess the effect of within-set similarity we computed a paired *t*-test between the repeatability scores obtained in Set 2 and Set 3.

2.4.2. Beer descriptions

To assess the effect of within-set similarity and choice of pivot, we analyzed the description of beers across sets or pivots. We selected the terms that best characterize each beer according to a hypergeometric law (Lebart, Piron, & Morineau, 2006). To evaluate the effect of within-set similarity, we compared the description of three beers common to two sets (Pelforth blond and Leffe blond common to Sets 1 and 2, Stella Artois common to Sets 2 and 3). To evaluate the effect of choice of pivot, we compared the three descriptions of the beers within Set 2 (obtained with the three pivots P1, P2, P3) and the two descriptions of the beers within Set 3 (obtained with P1, P2).

2.4.3. Beer sensory maps

To evaluate the effect of within-set similarity and choice of pivot, we compared the beer maps obtained via Correspondence Analysis[1] (CA) and Multiple Factor Analysis for Contingency Tables —MFACT— (Kostov, Bécue-Bertaut, & Husson, 2014) respectively. First, the negative and positive frequencies of each term were computed and the negative frequency subtracted from the positive one for each trial. The resulting score was then translated by adding the absolute value of the minimum score to all the scores (Thuillier et al., 2015). The minimum score thus takes on the value of zero and all other scores are positive. For each trial, the translated scores were finally compiled in a beers*terms matrix which was then submitted to CAs (Sets 1, 2 & 3 with P1) or to MFACTs (Set 2 with P1, P2, P3 and Set 3 with P1, P2).

3. Results

3.1. Effect of within-set beer similarity

To evaluate the effect of within-set beer similarity on PP outcomes we compared the data obtained for the three sets with P1 (blend) as Pivot (orange line in Fig. 1).

3.1.1. Performance criteria

Number of terms. There is an effect of within-set beer similarity on the number of terms [$F(2,20) = 3.707$; $p = 0.043$]. The more similar the beers in a set are, the smaller the number of terms used to describe them. Assessors used an average of 7.9 terms to describe Set 1 (low within-set similarity), of 7.1 terms for Set 2 (intermediate within-set similarity) and of 6.6 terms for Set 3 (high within-set similarity).

Consensus index. Globally, assessors are not very consensual, as showed by a consensus index close to 0 for the three sets. Moreover, although the ANOVA showed a significant set effect [$F(2,20) = 7.761$; $p = 0.003$], the consensus between assessors is not proportional to

[1] To evaluate the effect of within-set similarity, we performed CA analyses because it was not possible to performed a MFACT analysis as the beers are not the same in the three sets.

Table 2
Terms significantly more and less quoted by the assessors to describe the three beers common to two different sets with the P1 (blend) pivot. The sign + indicates that the beer was described as "more…" than the pivot and the sign − indicates that the beer was described as "less…" than the pivot.

		Set 1	Set 2
Pelforth Blond	More quoted	−Toasted −Alcoholic −Taste intensity +Cardboard −Fruity	+Toasted +Red fruits
	Less quoted	+Alcoholic +Fruity	+Fruity
		Set 1	Set 2
Leffe Blond	More quoted	+Phenol	+Phenol +Macerated fruits +Sweet +Odor intensity
	Less quoted	−Fruity	−Hoppy −Sweet
		Set 2	Set 3
Stella Artois	More quoted	+Bread +Yellow fruits +Malty −Sparkling	−Fruity
	Less quoted	NO TERM	NO TERM

the within-set beer similarity (Set 1, M = 0.076; Set 2, M = 0.025; Set 3, M = 0.070). The set effect might be due to the nature of the beers themselves rather than to their similarity.

Repeatability index. Globally assessors are not very repeatable as the repeatability index is equal to 0.127 for Set 2 and to 0.153 for Set 3. Moreover, no significant difference was found between the two sets ($t(20) = 0.258$; $p = 0.799$), indicating that repeatability is not influenced by within-set beer similarity.

3.1.2. Beer description

Table 2 presents the most and the least quoted terms for each beer common to two sets (Pelforth blond and Leffe blond common to Sets 1 and 2, Stella Artois common to Sets 2 and 3). Globally the beers common to two sets were not characterized with the same terms and the few similar terms are not used in the same direction (more/less) for the two sets. Pelforth blond was described as less *toasted* than the pivot in Set 1 and more *toasted* than the pivot in Set 2. Also, it was described as less *fruity* in Set 1 and more *red fruits* in Set 2. Leffe blond was described as more *phenol* than the pivot in both sets and Stella Artois more *yellow fruits* in Set 2 and less *fruity* in Set 3.

3.1.3. Beer sensory maps

The CA maps for Set 1, Set 2 and Set 3 are presented in Figs. 2–4 respectively. The first two factors of these CAs reach 67.8%, 66.1% and 57.7% of explained variance respectively indicating a similar variability of the data for the three sets. However, we can observe that the dispersion of the beer and term projections increases as a function of within-set beer similarity. The richness of the descriptions (*i.e.* number of contributing descriptors) and beer discriminability of the first two dimensions increase when the within-set beer similarity increases. The first dimension of Set 1 (low within-set similarity) opposes the terms *alcoholic* and *sweet* to *bitter* and *malty*; the second dimension opposes *sweet* and *malty* to *alcoholic* and *bitter* and separates the ChtiBL beer from the ChtiAMB. All the other beers and terms are plotted in the center of the map. The first CA dimension of Set 2 (intermediate within-set similarity) opposes the terms *sweet*, *phenolic*, *alcoholic* and *fruity* (Leffe and Grimbergen) to *malty (*Heineken1 and Stella) and the second dimension opposes *macerated fruits*, *burnt*, *hoppy*, *malty*, *phenolic* and *sweet* (Leffe) to *acidic*, *bitter*, *fruity*, *acidic persistence* and *bitter persistence* (Grimbergen). The first CA dimension of Set 3 (high within-set similarity) opposes the terms *floral*, *malty* and *cardboard* (Saint-Omer) to *alcoholic*, *bitter*, *hoppy*, *odor intensity*, *bitter persistence*, *fizzy* and *sweet* (Heineken1) and the second dimension opposes *acidic*, *alcoholic*, *bitter*, *malty* and *bitter persistence* (Heineken 2) to *fruity*, *sweet persistence*, *phenolic* and *sweet* (1664). It is interesting to notice that four terms (*alcoholic*, *sweet*, *bitter* and *malty*) are often used to describe the beers, whatever the set.

3.2. Effect of choice of pivot

To evaluate the effect of choice of pivot we compared the data from Set 2 with P1 (blend), P2 (Affligem) and P3 (Saint-Omer) as pivot on one side and from Set 3 with P1 (blend) and P2 (Jupiler) as a pivot on the other side (blue lines on Fig. 1).

3.2.1. Performance criteria

Number of terms. For both sets, there is no effect of choice of pivot on the number of terms used [For Set 2, $F(2,20) = 1.950$; $p = 0.168$ and for Set 3, $F(1,10) = 0.887$; $p = 0.368$]. For Set 2, assessors used an average of 7.1 terms with the blend pivot (P1), of 7.8 terms with the Affligem pivot (P2) and of 7.1 terms with the Saint-Omer pivot (P3). For Set 3, assessors used an average of 6.6 terms with the blend pivot (P1) and of 7.0 terms with the Jupiler pivot (P2).

Consensus index. Again, the assessors are not very consensual, as showed by a consensus index close to 0 for the two sets. Moreover, for both sets, the consensus between assessors is not influenced by the choice of pivot [For Set 2, $F(2,20) = 2.698$; $p = 0.092$ and for Set 3, $F(1,10) = 2.833$; $p = 0.123$]. For Set 2, the consensus indices for the blend pivot (P1), Affligem pivot (P2) and Saint-Omer pivot (P3) are 0.025, 0.066 and 0.058 respectively. For Set 3, the consensus indices for the blend pivot (P1) and for the Jupiler pivot (P2) are 0.070 and 0.101, respectively.

3.2.2. Beer description

Tables 3 and 4 present the most and the least quoted terms for each beer of Sets 2 and 3 with their different pivots. Globally for Set 2, the terms used to describe a given beer are quite different between the three pivots. Two exceptions are the Grimbergen which is described as *malty* with the three pivots and the Leffe blond which is described as *macerated fruits* with the three pivots and as *phenol* and *sweet* with two pivots. In the same way for Set 3, the descriptions of the beers are quite different from one pivot to another one, except for Heineken1 which is described as *banana* with the two pivots and St Omer which is globally described as a fruity beer with the P1 pivot (*fruity*) and the P2 pivot (*yellow fruits* and *citrus fruits*).

3.2.3. Beer sensory maps

Figs. 5 and 6 present the MFACT maps for Set 2 and Set 3 respectively. For Set 2, we can observe on the map 5b that the projections of the three tables are similar on the first dimension and that on the second dimension, the projection of table P1 (blend) is intermediate between the projections of table P2 and table P3. This observation is supported by the high values of RV coefficients between P1 and P2 (0.87) and between P1 and P3 (0.81). The RV coefficient between P2 and P3 is slightly lower (0.69). On the map 5a, we can observe that the pivot has more influence on the description of some beers than on others. For example, the influence of the three pivots on the Pelforth characterization is weak compared to the Leffe. P2 pivot (Affligem) tends to make Leffe more different from the other beers than P3 pivot (St Omer). The same can be observed for Grimbergen to a lesser extent. In contrast P3 pivot (St Omer) tends to differentiate more the Heineken from the other beers. For Set 3 (Fig. 6) we observe on map

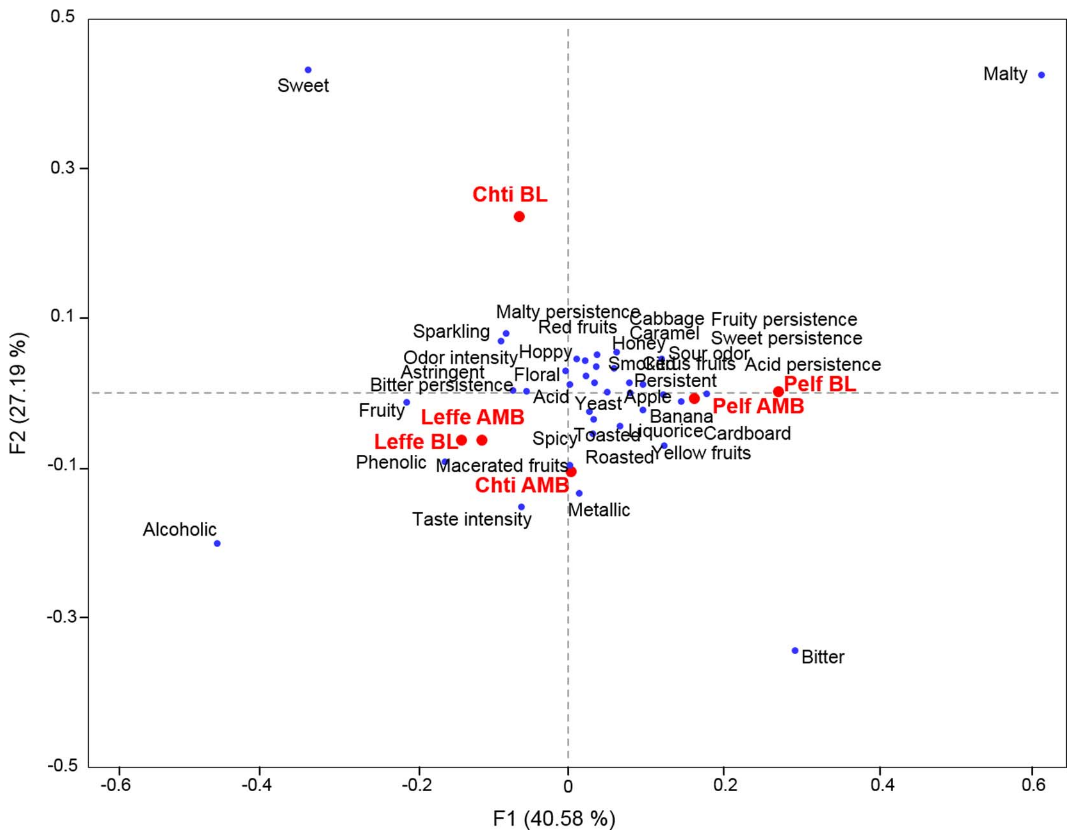

Fig. 2. CA map for the Set 1 with the P1 (blend) pivot.

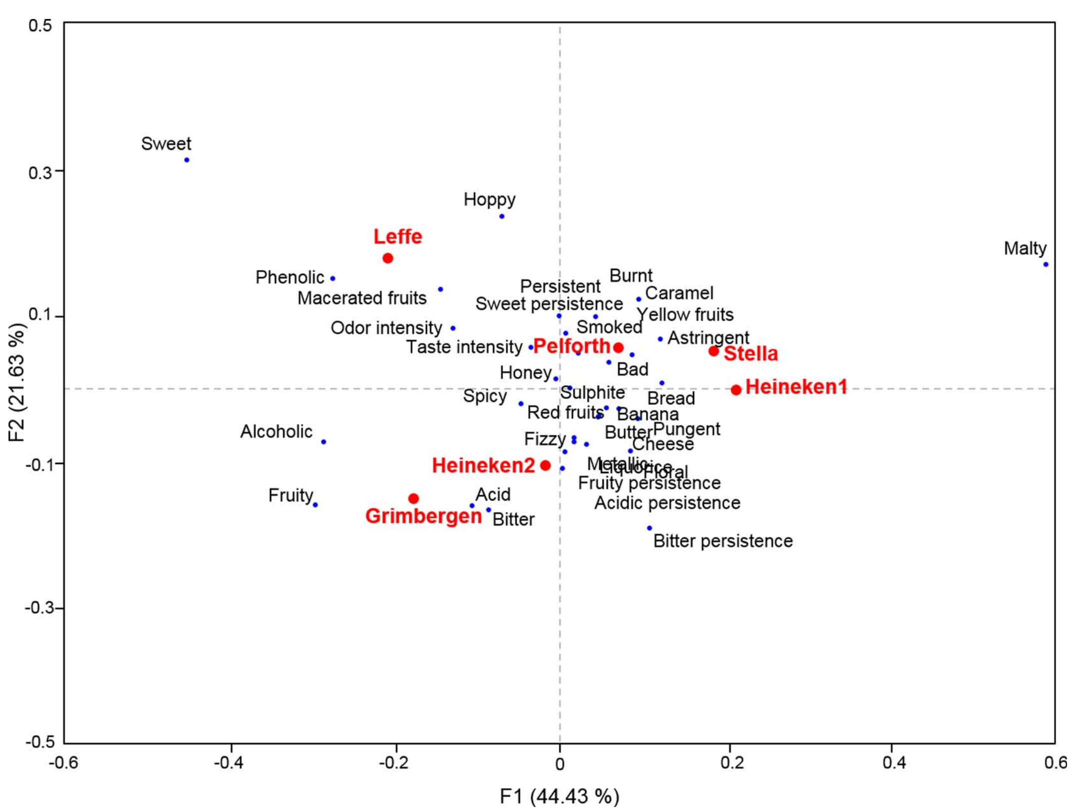

Fig. 3. CA map for the Set 2 with the P1 (blend) pivot.

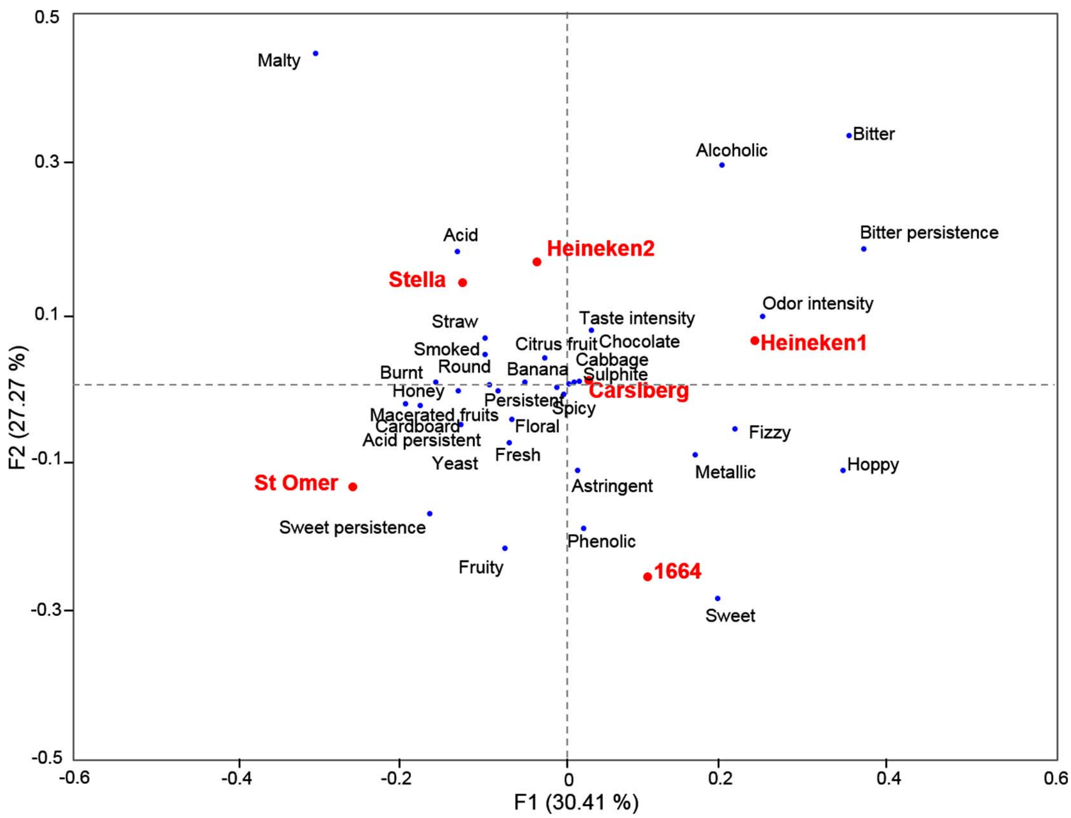

Fig. 4. CA map for the Set 3 with the P1 (blend) pivot.

Table 3
Terms significantly more and less quoted by the assessors to describe the beers of Set 2 with its three pivots. The sign + indicates that the beer was described as "more…" than the pivot and the sign − indicates that the beer was described as "less…" than the pivot.

		Blend Pivot (P1)	Affligem Pivot (P2)	Saint-Omer Pivot (P3)
Grimbergen	More quoted	+Fruity +Fruity persistence −Malty +Acidic persistence +Butter	+Cardboard	+Phenolic −Malty
	Less quoted	+Malty	NO TERM	+Bitter +Malty
Heineken 1	More quoted	−Persistent −Sweet −Alcoholic	+Bread	−Hoppy +Malty +Banana
	Less quoted	+Fruity	NO TERM	+Hoppy
Leffe blond	More quoted	+Phenolic +Macerated fruits +Sweet +Odor intensity	+Spicy +Macerated fruits +Phenolic +Fizzy	+Yellow fruits +Macerated fruits
	Less quoted	−Hoppy −Sweet	NO TERM	−Sweet
Pelforth blond	More quoted	+Burnt +Red fruits	NO TERM	+Bitter
	Less quoted	+Fruity	NO TERM	−Bitter +Odor intensity
Stella Artois	More quoted	+Bread +Yellow fruits +Malty −Fizzy	NO TERM	+Metallic −Fruity +Butter
	Less quoted	NO TERM	NO TERM	NO TERM

Table 4
Terms significantly more and less quoted by the assessors to describe the beers of Set 3 with the its two pivots. The sign + indicates that the beer was described as "more…" than the pivot and the sign − indicates that the beer was described as "less…" than the pivot.

		Blend Pivot (P1)	Jupiler Pivot (P2)
1664	More quoted	+Sweet +Phenolic	−Fizzy
	Less quoted	NO TERM	NO TERM
Carlsberg	More quoted	+Yeast +Cabbage	+Metallic +Persistent +Acidic
	Less quoted	+Fruity	NO TERM
Heineken 1	More quoted	+Banana	+Banana −Bitter persistence −Burnt
	Less quoted	NO TERM	NO TERM
St Omer	More quoted	−Bitter persistence −Odor intensity +Fruity	−Malty +Citrus fruits +Yellow fruits
	Less quoted	NO TERM	−Fizzy
Stella Artois	More quoted	−Fruity	+Alcoholic
	Less quoted	NO TERM	NO TERM

6b that the two tables are very close on both dimensions. The RV coefficient equal to 0.88 confirms this observation. This similarity between both tables is also observed on the map 6a and is quite the same for all the five beers.

4. Discussion

The objective of this paper was to evaluate a new descriptive method proposed by Thuillier et al. (2015). This new method relies on the comparison of a set of products with a stable reference —the pivot— and as such raises two main issues: the effect of within-set similarity and the choice of pivot. These two questions are discussed below.

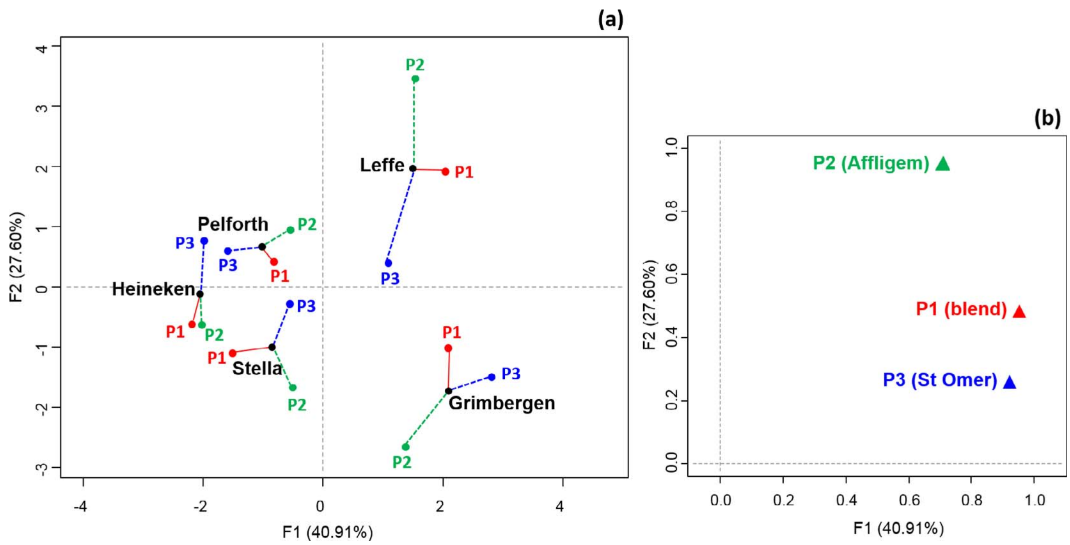

Fig. 5. MFACT maps for Set 2, (a) map of the beers with their respective projected points corresponding to their different pivots, (b) map of the three tables corresponding to the three pivots.

4.1. Influence of within-set beer similarity

We found a significant effect of within-set similarity on both the beer sensory maps and the vocabulary used to describe the beers. Surprisingly, the comparison of beer sensory maps showed that the discrimination between beers increased with the within-set beer similarity: The more similar the beers are the more the CA maps are scattered. Likewise, despite the fact that the total number of terms generated decreases with the within-set beer similarity, the richness of the descriptions (*i.e.* number of contributing descriptors) observed on CAs increase when beers are more similar. This result suggests that PP is more efficient for homogeneous spaces than for heterogeneous ones. The effect of within-set similarity was already reported in previous studies on alternative descriptive methods carried out with consumers. However, in these previous studies, more heterogeneous spaces led to better performance. For example, using three levels of sample complexity (defined as both the number of attributes used to describe the overall perception of the product and the degree of homogeneity of the sample set), Louw et al. (2014) observed a negative influence of this factor on the projective mapping results of alcoholic brandy products. Specifically, they reported that at high alcohol contents panelists' repeatability decreases when the complexity of the product set increases. Likewise, the ability of panelists to recognize duplicate samples (alcoholic beverages at 20% vol.) decreased with complexity. The authors interpret their results as an indication that evaluating a sample set with a relatively low degree of differentiation in the set is a more complicated task than evaluating a sample set with a high degree of differentiation. In agreement with this interpretation, Delarue and Sieffermann (2004) and Ares et al. (2015) found a lower similarity between the sensory maps obtained *via* a traditional descriptive analysis and those obtained *via* a Flash profile and a Check-all-that-Apply (CATA) method respectively, when the product spaces were more homogenous. Conversely PSP, another reference-based method, has shown good performance with very similar products, as mineral waters (Teillet, 2014; Teillet, Schlich, Urbano, COrdelle, & Guichard, 2010) or powered drinks (Antúnez et al., 2017) and even a better product discrimination than QDA when used with a trained panel (Varela, Svartebekk Myhrer, Naes, & Hersleth 2014). A plausible explanation for the fact that, contrary to other comparative methods (Flash profile, sorting task or projective mapping), PP is well adapted for restricted product spaces comes from the nature of the pivot. To evaluate the effect of within-set similarity we used a blend of all the samples of the

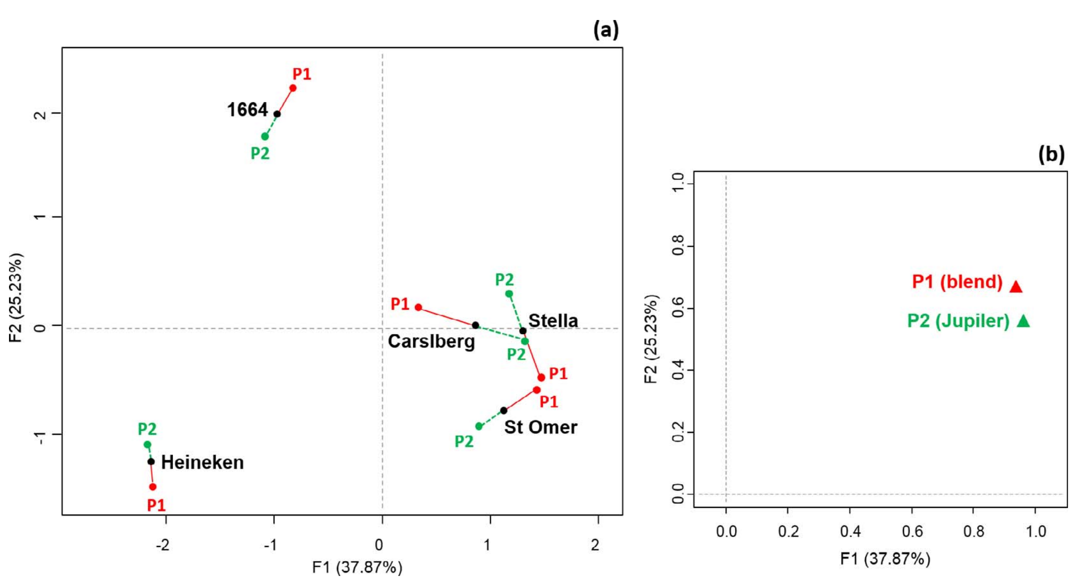

Fig. 6. MFACT maps for Set 3, (a) map of the beers with their respective projected points corresponding to their different pivots, (b) map of the two tables corresponding to the two pivots.

set as pivot. It is quite possible that this pivot is more representative of the whole set when the samples in this set are similar (*e.g.* Set 1) than when they are dissimilar (*e.g.* Set 3). And so the effect we observed might be due to the quality of the pivot rather than to the method itself. In other words, the idea of using a blend as pivot might be better adapted for homogeneous sets than for heterogeneous ones. Another explanation would be that with Set 3 (low within-set similarity), the blend pivot is very similar to each beer of the set, and so the assessors strove to find out the differences between the samples and the pivot and thus they went more into details than when the sensory differences are more obvious, as between the samples of Set 1 (high within-set similarity) and its blend pivot. This interpretation was already mentioned by Ares et al. (2015)) for PSP. However further work needs to be carried out to confirm and understand this result. Specifically, one way to validate the difference of product discrimination observed on the CA maps between the three sets would be to add confidence ellipses around the products. This would be possible by using bootstrapping approaches that consider individual variability (Cadoret & Husson, 2013) as it was already applied by Antúnez et al. (2017) to CA results and by Dehlholm, Brockhoff, and Bredie (2012) to MFA results. However, to do that a new script adapted to Pivot data needs no be developed, for example in the FactoMineR package of R. Our results highlight that further statistical development is necessary for the evaluation of the product discrimination and the stability of the perceptual spaces. This development was out of the scope of this paper.

Finally, we did not found an influence of the within-set similarity on the repeatability index calculated for the Heineken beers. A similar result was observed by Louw et al. (2014) on brandy products evaluated with projective mapping method. They found that at low alcohol levels (≈7% vol.), which are quite similar to the alcohol levels of our beers, the sample set complexity did not appear to affect panelists' repeatability whereas it is the case at high alcohol content (≈20% vol.). However, in Louw et al. (2014) the repeatability of the projective method was evaluated by presenting twice each sample set and then comparing the resulting configurations with RV coefficients and Relative Performance Indicator (RPI). In the present experiment, our repeatability index is based on the comparison of the terms used by each assessor to describe a repeated beer (Heineken). This index is very strict as it takes into account only the description of one repeated beer, which can explain the low values observed for both Sets 2 and 3. However, the analysis of the MFACT maps obtained for Sets 2 and 3 (Figs. 5 and 6 respectively) shows that the projections of the two repetitions of the Heineken beer (Heineken 1 and Heineken 2) are not very close, supporting the conclusions from the repeatability index. These results cannot be attributed to memory or attention issues as it could be the case for other alternative methods (*e.g.* sorting task, see Patris, Gufoni, Chollet, & Valentin, 2007). Indeed, in PP, participants evaluate the products by comparing them two by two. But it could be associated with saturation problems or sensory fatigue due to the large number of product pairs to be evaluated in one session, with such an alcoholic, bitter and persistent product as beer. The influence of the perceptual complexity of alcoholic products such as wine or beer on panel performance is also highlighted in the study of Louw et al. (2014). Other recent studies also evaluated the repeatability of different alternative descriptive methods. De Salamando et al. (2015) found that average individual repeatability was lower than global repeatability (evaluation of the same set of product two times by one group of assessors) for polarized projective mapping (PPM) and that individual reproducibility (evaluation of the same set of products by different groups of assessors) increased with the difference between the samples of the set. Also Hopfer and Heymann (2013) observed with projective mapping strong individual repeatability effects but not at the aggregate level of the consensus product maps. Antúnez et al. (2017) found that at the aggregate level, PSP, CATA and projective mapping on powdered drinks showed high repeatability even if some differences in the similarities and differences among the samples were identified between the three methods, especially for the narrower sample set. The repeatability of the PP method needs further research, for example by repeating the evaluation of an entire sample set.

4.2. Influence of the choice of pivot

In accordance with the results of Thuillier et al.'s simulations (2015), we found that the choice of pivot does not have a strong impact on the product positioning on the MFACT maps, as well as on the number of terms used to describe the beers and on the consensus between assessors. The suggestion of Thuillier et al. to create a "central product" by blending all products to be described seems to be a good option when it is possible (mixable products). Indeed, we observed on the MFACT of Set 2 that the blend pivot (P1) has a central position between the two other extreme pivots (P2 and P3), indicating that the product descriptions obtained with the blend pivot are intermediate between the ones obtained with the extreme pivots. These results also showed that no physicochemical interactions occurred with the blending of the beers that modified the pivot beer sensory characteristics in a radical way. In the case of non-mixable products, the choice of the pivot should not be a big issue as long as the pivot belongs to the product space (*e.g.* dark beer pivot for dark beer space) since quite similar results were obtained with the different tested pivots. PSP is another relatively new method based on the comparison of samples to reference products, usually called poles (Teillet et al., 2010). The influence of the choice of the poles on PSP results has been also questioned. If small changes in the sets of poles do not lead to relevant changes in sample configurations (De Salamando et al., 2015; Teillet, 2014), Ares et al. (2015) insisted on the fact that the poles should really reflect the main characteristics responsible for the expected similarities and differences among the samples to be evaluated. So compared to PSP, the choice of pivot in PP does not seem to be a big issue. However, this statement needs to be verified by directly comparing PSP and PP on the same set of samples, changing the poles/pivots and looking at the stability of the perceptual spaces. This difference could be explained by the nature of the task. In PSP, even if no indication is given to the consumers on the sensory attributes to use for comparing the samples to the poles, they tend to focus on one or two characteristics, even more when the poles are strongly associated with these specific characteristics (Ares et al., 2015). Conversely in PP the assessors do not limit their evaluation to few sensory attributes, as shown by the average number of terms used to describe each beer (≈7 terms). Again, it would be interesting to compare these two reference-based methods to confirm this hypothesis in order to gain insight for more relevant applications.

Yet if the choice of pivot does not affect directly the product sensory maps, the sensory quality of the pivot affects the product descriptions. This appeared clearly when we compare the descriptions of the beers common to two sets (Pelforth blond and Leffe blond for Sets 1 and 2; Stella Artois for Sets 2 and 3): The same beer is not described with the same words depending on the set it belongs to. This result is quite logical as the common beers are described in comparison to a blend pivot (P1), which is not the same from one set to another. Because the products are not described in the absolute but by reference to the pivot, the terms generated for a given product are in part dependent on the sensory characteristics of the pivot. Let us consider a very bitter pivot beer. If the sample is not bitter at all, it is likely that the assessors will use the term *bitter* to describe the difference between the pivot and the sample. Conversely if the sample is also very bitter, the term *bitter* will probably not be used to characterize the difference between the pivot and the sample. So depending on the sensory characteristics of the pivot, the terms used to describe the samples will not be the same even though the relative positioning of the samples will be similar whatever the pivot. This result suggests that PP would be a good methodological choice for obtaining global information about the similarities and differences between samples by analyzing multidimensional maps but is less appropriate for their detailed sensory characterizations. In this way, PP is more similar to similarity-based methods (*e.g.* sorting task, projective mapping) than to attribute-based methodologies (*e.g.* CATA, free choice profiling) which encourage the assessors to focus their attention on the sensory characteristics and so provide more detailed descriptions (Varela & Ares, 2012). Besides, PP is useful when comparison with a reference product is of particular interest as well as when all the samples cannot be tested in one session.

5. Conclusion

Pivot Profile© (PP) is a fast descriptive method proposed recently by Thuillier et al. (2015) as an alternative to be carried out with professionals who often prefer using free descriptions than classical descriptive methods. Our study suggests that the choice of pivot has less influence than the within-set similarity (similarity within the product space) on the beer positioning on the sensory maps. It seems that the PP method is more suitable for restricted product spaces in terms of sensory characteristics, and that the creation of a "central product" as the pivot can be a good option when the type of products allows it. It also suggests that PP would be a good method for obtaining the similarities and differences between the products but not to access to the detailed sensory profiling of each product.

Other points are raised by this study and need further research. The first one concerns PP performance in terms of repeatability and consensus between assessors. The indices used to evaluate these criteria were probably too strict as they evaluate individual repeatability only. Evaluating both individual and group repeatability as in QDA might be a better approach and further development in this direction are needed. Also, the number of assessors can be questioned. Even if we used highly trained panelists who are more consensual and repeatable than consumers, 11 assessors might not be enough to obtain reliable PP results. One way to address this issue would be to apply a bootstrapping resampling technique to evaluate the minimum number of assessors necessary to obtain a stable product space, as suggested by Blancher, Clavier, Egoroff, Duineveld, and Parcon (2012). This would require developing a new script adapted to the PP as was done for sorting task with Indscal (Bárcenas, Pérez Elortondo, & Albisu, 2004; Nestrud & Lawless, 2011) and Distatis (Abdi, Valentin, Chollet, & Chrea, 2007; Abdi, Valentin, O'Toole, & Edelman, 2005).

Although further work is needed in this way, we believe that the number and the nature of assessors needed to perform PP varies according to the objective of the task. If we just want a coarse description of the products obtained from CA map, a few assessors might be enough. However, if the objective is to obtain precise descriptions, having more assessors might be necessary. As suggested by Thuillier et al. (2015), PP seems to be well suited for expert panelists but it is sometime difficult to find a large number of expert assessors. So it would be important to check whether PP could also be used with consumers.

References

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

Abdi, H., Valentin, D., Chollet, S., & Chrea, C. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and application. *Food Quality and Preference*, *18*, 627–640.

Antúnez, L., Vidal, L., de Salamando, L., Giménez, A., & Ares, G. (2017). Comparison of consumer-based methodologies for sensory characterization: Case study with four sample sets of powered drinks. *Food Quality and Preference*, *56*, 149–163.

Ares, G., Antúnez, L., Bruzzone, F., Vidal, L., Pineau, B., Beresford, M. K., et al. (2015). Comparison of sensory product profiles generated by trained assessors and consumers using CATA questions: Four case studies with complex and/or similar samples. *Food Quality and Preference*, *45*, 75–86.

Ares, G., Antúnez, L., Oliveira, D., Alcaire, F., Berget, I., Næs, T., et al. (2015). Pole selection in Polarized Sensory Positioning: Insights from the cognitive aspects behind the task. *Food Quality and Preference*, *46*, 48–57.

Bárcenas, P., Pérez Elortondo, F. J., & Albisu, M. (2004). Projective mapping in sensory analysis of ewes milk cheeses: A study on consumers and trained panel performance. *Food Research International*, *37*, 723–729.

Blancher, G., Clavier, B., Egoroff, C., Duineveld, K., & Parcon, J. (2012). A method to investigate the stability of a sorting map. *Food Quality and Preference*, *23*, 36–43.

Cadoret, M., & Husson, F. (2013). Construction and evaluation of confidence ellipses applied at sensory data. *Food Quality and Preference*, *28*, 106–115.

Chollet, S., Valentin, D., & Abdi, H. (2014). Free sorting task. In P. Varela & Arès, G. (Eds.), Novel techniques in sensory characterization and consume profiling (pp. 207–227). Boca Raton: CRC Press.

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.

De Salamando, L., Antúnez, L., Torres-Moreno, M., Giménez, A., & Ares, G. (2015). Reliability of polarized projective mapping with consumers. *Journal of Sensory Studies*, *30*, 280–294.

Dehlholm, C. (2014). Projective mapping and Napping. In P. Varela & Arès, G. (Eds.), Novel techniques in sensory characterization and consume profiling (pp. 229–253). Boca Raton: CRC Press.

Dehlholm, C., Brockhoff, P. B., & Bredie, 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.

Delarue, J. (2014). Flash Profile. In P. Varela & Arès, G. (Eds.), Novel techniques in sensory characterization and consume profiling (pp. 175–205). Boca Raton: CRC Press.

Delarue, J., & Sieffermann, J.-M. (2004). Sensory mapping using Flash profile. Comparison with a conventional descriptive method fort the evaluation of the flavor of fruit dairy products. *Food Quality and Preference*, *15*, 383–392.

Fonseca, F. G. A., Esmerino, E. A., Tavares Filho, E. R., Ferraz, J. P., da Cruz, A. G., & Bolini, H. M. A. (2016). Novel and successful free comments method for sensory characterization of chocolate ice cream: A comparative study between pivot profile and comment analysis. *Journal of Diary Science*, *99*, 3408–3420.

Hopfer, H., & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference*, *28*, 164–181.

Kostov, B., Bécue-Bertaut, M., & Husson, F. (2014). An original methodology for the analysis and interpretation of word-count based methods: Multiple factor analysis for contingency tables complemented by consensual words. *Food Quality and Preference*, *32*, 35–40.

Lebart, L., Piron, M., & Morineau, A. (2006). *Statistique exploratoire multidimensionnelle. Visualisation et inférence en fouille de données*. Paris: Dunod.

Louw, L., Oelofse, S., Næs, T., Lambretchts, M., van Rensburg, P., & Nieuwoudt, H. (2014). Trained sensory panelists' response to product alcohol content in the projective mapping task: Observations on alcohol content, products complexity and prior knowledge. *Food Quality and Preference*, *34*, 37–44.

Nestrud, M. A., & Lawless, H. T. (2011). Recovery of subsampled dimensions and configurations derived from napping data by MFA and MDS. *Attention, Perception, & Psychophysics*, *73*, 1266–1278.

Patris, B., Gufoni, V., Chollet, S. & Valentin, D. (2007). Impact of training on strategies to realize a beer sorting task: behavioral and verbal assessments (pp. 17–29). In D. Valentin, D.Z. Nguyen & L. Pelletier (Eds.) New trends in Sensory Evaluation of Food and Non-Food Products. Ho Chi Minh, Vietnam: Vietnam National University-Ho chi Minh City Publishing House.

Risvik, E., McEwan, J. A., & Redbottena, M. (1997). Evaluation of sensory profiling and projective mapping. *Food Quality and Preference*, *8*, 63–71.

Teillet, E. (2014). Polarized Sensory Positioning Methodologies. In P. Varela & Arès, G. (Eds.), Novel techniques in sensory characterization and consume profiling (pp. 255–270). Boca Raton: CRC Press.

Teillet, E., Schlich, P., Urbano, C., COrdelle, S., & Guichard, E. (2010). Sensory methodologies and the taste of water. *Food Quality and Preference*, *21*(8), 967–976.

Thuillier, B., Valentin, D., Marchal, R., & Dacremont, C. (2015). Pivot profile: A new description method based on free description. *Food Quality and Preference*, *42*, 66–77.

Valentin, D., Chollet, S., Lelièvre, M., & 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*(8), 1563–1578.

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., Svartebekk Myhrer, K., Naes, T., & Hersleth, M. (2014). The best of both worlds: rapidity and enhanced discrimination with a trained panel. Exploration of global and partial PSP with a descriptive step. In 6th European Conference on Sensory and Consumer Research, 7–10 September 2014, Copenhagen, Denmark.