Food Quality and Preference 48 (2016) 41–49

Contents lists available at ScienceDirect

Food Quality and Preference

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

Performance of Flash Profile and Napping with and without training for describing small sensory differences in a model wine

Jing Liu [a], Marlene Schou Grønbeck [a,1], Rossella Di Monaco [b], Davide Giacalone [a,2], Wender L.P. Bredie [a,*]

[a] Department of Food Science, Faculty of Science, University of Copenhagen, Rolighedsvej 30, DK-1958 Frederiksberg C, Denmark
[b] Food Science and Agricultural Department, University of Naples-Federico II, Via Università 100, P.co Gussone, 80055 Portici, Naples, Italy

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form 18 August 2015
Accepted 20 August 2015
Available online 20 August 2015

Keywords:
Napping
Flash Profile
Rapid sensory methodologies
Sensory analysis
Wine
Training/modification

ABSTRACT

Rapid sensory methods are a convenient alternative to conventional descriptive analysis suitable for quickly assessing sensory product differences. As these methods gain in popularity, assessments of their discriminability and reproducibility in food applications are increasingly needed. Moreover, it is of interest to explore whether small adjustments to the existing protocols could improve the results. In this study different variations of two rapid sensory methods, one based on holistic assessment – Napping, and one based on attribute evaluation – Flash Profile, were tested for the evaluation of the flavour in wine. Model wines were developed with control over the sensory differences in terms of sensory characters and sensory intensities (weak to moderate). Some modifications to the classical Napping and Flash Profile protocols were employed in order to improve discriminability, repeatability and accuracy. The results showed that conducting Napping with a panel training on either the method (training on how to arrange samples on the sheet) or the product (familiarisation with the sensory properties of the wines) improved the outcome compared to the classical Napping protocol. The classical Flash Profile protocol and its modified version including a Napping with subsequent attributes generation as the word generation step and limiting the number of attributes for ranking gave a similar sample space. The Napping method could best highlight qualitative sample differences, whereas the Flash Profile provided a more precise product mapping on quantitative differences between model wines.

1. Introduction

Rapid sensory methods have gained considerable interests as alternatives to conventional descriptive profiling, due to their speed and cost-effectiveness. Among the several alternatives proposed, Napping and Flash Profile are two attractive approaches. Napping (Pagès, 2003) is a specific variant of Projective Mapping, a method originally proposed for applied sensory studies by Risvik, McEwan, Colwill, Rogers, and Lyon (1994) to describe overall differences among samples. The samples are simultaneously presented to the assessors, who are then required to project samples on a two-dimensional space in a way that reflect their perceived sample differences, i.e., by placing samples perceived as similar close to each other, and samples perceived to be more different further apart. Since Napping itself does not provide a description of the samples, a subsequent step where the assessors write down attributes to describe samples is usually coupled with Napping. Although many researchers have reported that Napping is an easy and user-friendly method to use (Albert, Varela, Salvador, Hough, & Fiszman, 2011; Veinand, Godefroy, Adam, & Delarue, 2011), it has been found that, without a proper training on the method, some assessors may have problems with the Napping task. For instance, they might be unable to create a plane sample representation map (Hopfer & Heymann, 2013; Nestrud & Lawless, 2008; Pagès, 2005; Veinand et al., 2011). In order to overcome this limitation efforts have been made to conduct panel training before performing the Napping task, using different approaches. Risvik, Barcenas and their colleagues (Barcenas, Elortondo, & Albisu, 2004; Risvik, McEwan, & Rødbotten, 1997; Risvik et al., 1994) used the example of intercity distances from Kruskal and Wish (1978), and Hopfer and Heymann (2013) took a short training exercise using various shapes differing in colour and size. Besides, earlier works on training of sensory panels for descriptive analysis showed rapid learning in the early confrontations of the panel with the products (Byrne,

* Corresponding author at: Department of Food Science, Section for Sensory and Consumer Science, University of Copenhagen, Rolighedsvej 30, 1958 Frederiksberg C, Denmark.

E-mail address: wb@food.ku.dk (W.L.P. Bredie).

[1] Current address: Department of Food Science, Aarhus University, Kirstinebjergvej 10, 5792 Årslev, Denmark.

[2] Current address: Department of Technology and Innovation, University of Southern Denmark, Campusvej 55, 5230 Odense, Denmark.

http://dx.doi.org/10.1016/j.foodqual.2015.08.008
0950-3293/

Bredie, & Martens, 1999; Byrne, O'Sullivan, Dijksterhuis, Bredie, & Martens, 2001; Liu et al., 2015). It is therefore suggested to expand the classical Napping protocol (and other perceptual mapping methods) with a simple panel product orientation session. To date, no study has yet investigated the outcome from different training strategies on Napping compared with the classical Napping approach.

Flash Profile (FP) was introduced by Sieffermann (2000) as a variant of Free Choice Profiling (Williams & Langron, 1984), in which the assessors are required to have a comparative assessment of the whole sample set. Assessors are asked to list the sensory characteristics that best describe the differences among the samples and then rank all the samples for each of their individual attribute lists (Dairou & Sieffermann, 2002). FP has already been used for sensory evaluation in many different food product categories, including jams (Dairou & Sieffermann, 2002), dairy products (Delarue & Sieffermann, 2004), hot beverages (Moussaoui & Varela, 2010), lemon iced teas (Veinand et al., 2011), fish nuggets (Albert et al., 2011), and liver pâté (Dehlholm, Brockhoff, Meinert, Aaslyng, & Bredie, 2012). One of the drawbacks of this method is that, since assessors are not imposed on the use of a common vocabulary, the semantic interpretation of FP results can be complex. Besides, if assessors generate a large amount of attributes, the difficulty of ranking samples would increase significantly. The attributes generation step in FP thus plays an important role in the quality of the results obtained with this method. Perrin and Pagès (2009) hypothesised that attributes collected in the context of a holistic sensory task (i.e., Napping or Free Sorting Task) could be more consistent than those obtained in a totally free semantic description. They pointed out that the positioning task could help assessors to get familiar with the sample space. Inspired by this idea, it might be interesting to include Napping as a product familiarisation and attributes generation step within the FP method. Perrin et al. (2008) had already applied a similar approach on white wines, where attributes from the Napping tablecloth were collected prior to running a Free Choice Profiling. In the context of the FP method, this might be a valuable modification to increase assessors' focus on the most important attributes within the sample set. In the present study, we tried to test the outcome of this modification by comparing a classical FP and its modified version.

It should be noted here that, as one of holistic approaches and attribute-based methodologies, Napping and FP have been compared in previous studies using different food products. For instance, Albert et al. (2011) performed a Napping, FP and conventional descriptive analysis on hot served food fish nuggets. They found that FP provided more detailed information about the samples characteristics while Napping tended to summarise the information. Moussaoui and Varela (2010) conducted a Sorting, Projective Mapping, FP and Repertory Grid Method (RGM) on hot beverages, reporting that FP and RGM presented the advantage of producing more relevant and richer descriptions comparing to the holistic methodologies. Dehlholm, Brockhoff, Meinert, et al. (2012) compared Napping, FP and several other methodologies on liver pâtés, and observed that FP had higher discriminability than Napping. However, a limitation of such studies is that most have been carried out on complex food products. Usually, conventional descriptive analysis was used as a benchmark to infer validity, but actual differences among products were unknown. In the present work, we chose to use a model system developed with full control over the sensory differences among samples, both in terms of sensory characters and perceived intensities, in order to more accurately evaluate the performance of the different methodologies.

A last issue to be considered is evaluating the repeatability and accuracy of rapid methodologies. Previous works have been done on this topic by comparing responses from the same assessors to the same sample set in different sessions for the repeatability, or comparing the position of a blind repeated sample on product spaces for accuracy (Ares & Varela, 2014, chap. 14). Several researchers have found that both Napping and FP were repeatable (Dairou & Sieffermann, 2002; Moussaoui & Varela, 2010; Veinand et al., 2011). The FP method has normally been performed in two or three replicates (Dairou & Sieffermann, 2002; Price et al., 2014). On the contrary, it is less prevalent to include replicates for the Napping, and also the use of blind duplicate within the same session has not been extensively used to evaluate this method. Large differences from replicate to replicate in the sample configurations of Napping have been observed in previous studies although overall similarities and differences among samples were constant over repeated sessions (Ares & Varela, 2014, chap. 14; Hopfer & Heymann, 2013; Kennedy, 2010; Mielby, Hopfer, Jensen, Thybo, & Heymann, 2014). Therefore, in this study we looked into the effects of replicate between and within sessions on each of used method.

Overall, the specific objectives of this study were:

(a) to evaluate whether conducting training prior to Napping could enhance the outcome, and to evaluate which type of training (method or product) would give the best results;
(b) to investigate whether a modification could improve the results obtained with the Flash Profile approach;
(c) to compare the performance of Napping and Flash Profile on wine samples with small and experimentally controlled differences;
(d) to evaluate the repeatability of the methodologies between sessions (using measures of configurational similarity) and the accuracy within sessions (using the position of a blind repeated sample).

2. Materials and methods

2.1. Samples

2.1.1. Sample preparation

Model wines were prepared by adding either one of the flavour compounds benzaldehyde, isopentyl acetate or 2-phenylethanol into a plain white wine (Pinot Blanc, 11.5% v/v, Alsace) at two concentrations labelled as 'high' and 'low'. The flavour compounds were of food grade quality and of a high purity (Aldrich, USA). The concentrations were decided based on a series of pre-tests, and verified by a scaling test (Section 2.1.2) in order to make sure the wine samples had small but detectable differences. Two samples – 'benzaldehyde high' and '2-phenylethanol low' – were served as blind duplicates to investigate the performance of the panel. So nine wine samples in total, including a base wine without any addition of flavours, were used for the sensory tests (Section 2.2). The details on the model wine samples are shown in Table 1.

2.1.2. Samples verification

In order to verify the sensory differences between the model wines, a scaling test on the perceived intensities was performed with a sensory panel of ten assessors (three males; mean age = 32 years). This formed a basis for comparison of the Napping and FP evaluations. The scaling test was run in four replicates, in each of which the assessors were asked to rate the intensity of four attributes (almond, banana, rose/floral/spicy, and overall intensity) on unstructured 15-cm scales for each of the seven wine samples (two blind replicates were not tested in this section).

Table 1
List and details of model wine samples used in the study.

Flavour compounds added	Level of addition (μL/L)	Odour descriptors[b]
Benzaldehyde High[a]	8	Almond
Benzaldehyde Low	2	Almond
Isopentyl acetate High	5	Banana
Isopentyl acetate Low	1	Banana
2-Phenylethanol High	100	Rose/floral/spicy
2-Phenylethanol Low[a]	10	Rose/floral/spicy
Base wine	–	A plain white wine

[a] The two wines were used as blind replicates.
[b] Odour descriptors reported in Flavournet (http://www.flavornet.org/f_kovats.html).

2.2. Experimental procedures

Three different approaches of Napping and two approaches of FP were performed. An overview of all the rapid sensory tests is shown in Fig. 1. The assessors were asked to focus only on the flavour in each test. They were instructed to rinse with water and/or a cracker between samples to minimise carry over effects such as adaptation. They were also instructed to spit out the samples. Each test was carried out in duplicate, and each repeated session was done on separate days within one week. In all the sensory tests, the samples were served as 25 mL aliquots at 10 ± 1 °C in standardised wineglasses (ISO-3591, 1977), which were coded with 3-digit numbers and covered with watch glasses.

All the sensory evaluations took place in the sensory laboratory at the University of Copenhagen, Denmark.

2.2.1. Assessors

The assessors were recruited from the sensory panel at the University of Copenhagen. All were generally trained in sensory evaluation and had participated in previous studies of different food matrices, including wine. Four different panels were recruited for performing the five different sensory approaches, each consisting of eight or nine assessors. The panel which did the classical Napping was asked to run the modified FP later since the Napping was part of the modified FP. The assessors who attended the scaling test were allowed to be part of these four panels due to limited assessors, but they came back for the rapid tests after at least one month so as to prevent learning effects. Also, they were equally distributed into these four panels. All assessors were compensated for their participation.

2.2.2. Napping procedures

Three different approaches of Napping were performed, one (A) based on the classical Napping protocol proposed by Pagès (2003) and two others (B and C) including training sessions prior to the Napping (Fig. 1). The assessors received the training before the first round of Napping, with no further training prior to the second replicate. For the Napping panels, all the nine wine samples (including two blind replicates) were simultaneously presented to each assessor. They were requested to lay out the wine samples on an A2 paper (~60 × 40 cm) in such a way that two wine samples were placed very close to each other if they seemed identical and that two wines were distant from one another if they seemed different. Assessors then added attributes directly on the sheet to describe the sensory differences and similarities of the wines.

(A) Classical Napping

A Napping with no training phase was carried out by a panel of nine assessors (panel A: two males, mean age = 34 years). An oral and printed description of the task was provided to the assessors.

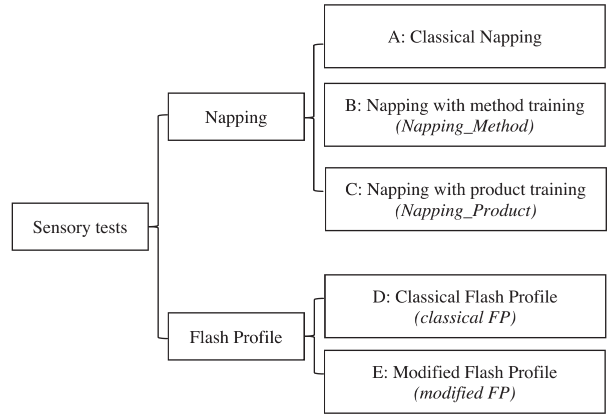

Fig. 1. Overview of experimental procedures. Four different panels were recruited for five different approaches. The panel employed in the classical Napping (panel A) was later used for the modified FP ranking evaluation.

'*Please evaluate the wine samples and place them on the A2 paper according to how similar or dissimilar they are to you. The more similar the wines are the closer they should be positioned to each other, the more dissimilar they are the further apart they should be positioned. There is no wrong way to do this! You are welcome to use the whole paper. After placing all the samples on paper, please write the 3-digit numbers on the paper. Then write down a few descriptive words below the 3-digit numbers to describe several samples if you want. Finally, please write your name on the paper. Press the switch, if you have any questions. Thank you for your help.*'

(B) Napping with method training (*Napping_Method*)

A different panel of nine assessors (panel B: two males, mean age = 28 years) received a training exercise on the Napping method before they participated in the test. The training session was conducted as described by Hopfer and Heymann (2013), where the assessors were given different paper shapes that differed in colour, shape and size. They were asked to place the figures on an A2 paper using their own criteria and write down some words describing these criteria. The panel leader observed the assessors' way of performing and guided them in using the whole paper and understanding the concept. The assessors were not corrected or told to follow a standardised way. All the assessors received this training individually so that they would not influence each other. The aim of the training session was to guide the assessors to understand the Napping task.

(C) Napping with product training (*Napping_Product*)

A training session on the products was applied to a third panel of nine assessors (panel C: four males, mean age = 27 years). The training was divided into four different steps: testing references, recognising references, comparison test on flavour intensities and comparison test on flavour characters. The first step was providing the references to the panel. A list of the references with concentrations and descriptions can be seen in Table 2a. The second step was to see whether the assessors were able to recognise the references. Three glasses were provided, each with the high concentration used in the first two Napping tasks (see Table 1). In the last two steps, two sets of samples with the same added flavour compound but in the low and high concentration were served to train the assessors on intensity differences and another two sets of samples with different flavour compounds but with the same intensity level were presented to train the assessors on character differences. The wine samples used for step 3 and 4 are shown in Table 2b. After a

Table 2a
The wine samples used in the training phase step 1 in the Napping with product training.

Flavour compounds added	Level of addition (μL/L)	Odour descriptors
Benzaldehyde High	16	Almond
Isopentyl acetate High	10	Banana
2-Phenylethanol High	200	Rose/floral/spicy
Base wine	–	A plain white wine

20-min break the panel conducted the Napping task on the nine wine samples.

2.2.3. Flash Profile procedures

Two different FP protocols were performed, a classical FP (D) and a FP proceeded with Napping and individual vocabulary restrictions (E), labelled '*modified FP*' in Fig. 1 and in the remainder of the paper.

(D) Classical FP

A classical FP was carried out by a fourth panel of eight assessors (panel D: three males, mean age = 29 years) according to the protocol described in Dairou and Sieffermann (2002). It consisted of three sessions in which the nine samples were presented simultaneously. In the first session, the assessors were given an explanation about the procedure, and then were asked to individually list the sensory characteristics that best described the differences among the samples. They were instructed to avoid the use of hedonic terms. The attribute generation process lasted around 30 min. After a short break the assessors were provided a pooled list of attributes derived from all assessors on a white board. They were allowed to modify their lists by adding or deleting any of their own attributes. In the final two sessions, assessors were requested to rank the wine samples on their own attributes according to intensity differences. They placed the codes of the samples on a line scale anchored at the left side 'low intensity' and right side 'high intensity'. Ties were allowed.

(E) Modified FP

The same panel employed in the classical Napping (panel A) was later used for the modified FP ranking evaluation, except one assessor who was absent, so the panel consisted of eight assessors (two males, mean age = 30 years). The modified FP included two additions to the classical FP approach. Firstly, a Napping followed by an attribute generation step was integrated as a way to help assessors get familiar with the product space, as well as focus on the attributes that were discriminating the samples. The second modification was that assessors were restricted to use maximum 10 relevant attributes for FP ranking, with the expectation that reducing the number of potential attributes would help the assessors focus on the core differences perceived from the samples and thus decrease the noise level in the data analysis.

2.3. Data analysis

2.3.1. Scaling data for samples verification

An analysis of variance (ANOVA) followed by the least significant difference (LSD) post hoc test was carried out using IBM SPSS Statistics 22 (IBM Corporation, Armonk, NY). Differences were considered significant if the *p*-value was below 0.05.

2.3.2. Napping and FP data

For the Napping data, the *X* and *Y* coordinates on the paper were measured for each sample relative to the left bottom corner of the sheet. The data in each Napping were then organised into 9 rows representing the samples, in 2×9 (where 9 is the number of assessors) columns. Since all Napping evaluations were performed in duplicate with nine assessors, the final data matrix consisted of 9 rows $\times$ 36 columns ($2 \times 9 \times 2$) for the multivariate analyses.

To analyse the FP data, individual matrices for each assessor (samples $\times$ attributes) were assembled. As the evaluation was performed as rank ordering, each sample consisted of the rank value from 1 to 9 given to that sample. For samples with equal rank orders (tied samples) a mean value was used. The FP data set was grouped by assessor and averaged over individual repetitions.

Both Napping and FP data were analysed by Multiple Factor Analysis (MFA). After running the MFA, bootstrapping was performed on the factor scores of the samples, in order to obtain confidence ellipses accounting for 95% of the bootstrapped values (Dehlholm, Brockhoff, & Bredie, 2012). The 95% confidence ellipses were applied to graphically display uncertainty within individual methodological approaches and to compare differences in uncertainty between methods. The RV coefficient (Escoufier, 1973) was used to compare the overall sample configurations between replicates within each individual method. The RV coefficient takes values between 0 and 1, with values closer to 1 indicating higher similarity. The significance of the coefficients computed was evaluated using permutation tests as suggested by Josse, Pagès, and Husson (2008). All multivariate analyses were carried out in R version 2.12.1 (R Development Core Team, 2010) applying functions from the package FactoMineR (Lê, Josse, & Husson, 2008) and SensoMineR (Lê & Husson, 2008).

3. Results

3.1. Samples verification

The ANOVA results for the scaling data are presented in Table 3. There were significant differences at $p < 0.001$ for benzaldehyde and isopentyl acetate sample sets with different concentrations; for 2-phenylethanol the difference was significant at the $p < 0.05$ level, indicating that 2-phenylethanol was the most difficult one for the assessors to distinguish between different concentrations. The result confirmed that the sensory differences in the intensities of the model wines could be detected, but some of these were relatively small. This was however consistent with the intended design aiming at creating a sample space with small sensory differences in order to better evaluating the performance of Napping and FP.

Table 2b
The wine samples used in the training phase step 3 and 4 in the Napping with product training.

Steps	Tests	Sample 1	Level of addition (μL/L)	Sample 2	Level of addition (μL/L)
Step 3	1 – Intensity Test	Benzaldehyde Low	2	Benzaldehyde High	8
Step 3	2 – Intensity Test	2-Phenylethanol Low	10	2-Phenylethanol High	100
Step 4	3 – Flavour Test	Isopentyl acetate Low	1	2-Phenylethanol Low	10
Step 4	4 – Flavour Test	Benzaldehyde High	8	2-Phenylethanol High	100

Table 3
The ANOVA results for the scaling data. Mean values for each sample are shown in the table; values with different superscript letters are significant. $^{***}p < 0.001$; $^{*}p < 0.05$.

Flavours	Compounds	Base	Low	High	*F* values	Significance
Almond	Benzaldehyde	3.48^a	6.16^b	11.88^c	19.74	***
Banana	Isopentyl acetate	3.29^a	5.25^b	10.81^c	24.11	***
Rose/floral/spicy	2-Phenylethanol	5.59^a	6.47^b	7.69^c	2.71	*

3.2. Napping

3.2.1. Comparison of configurations obtained with the three Napping approaches

The scores plots with 95% confidence ellipses obtained by MFA from the Napping data are displayed in Fig. 2. In the classical Napping, the first two dimensions accounted for 44% of the explained variance (24% and 20%, respectively). An increase of the total explained variability was observed in Napping with training strategies. Compared to the classical Napping (44%), the total variance of the *Napping_Method* increased to 56%; the *Napping_Product* showed a much larger increase with an overall 67% variance explained in the first two dimensions. The training either on method or product increased the explained variance, and thus reduced the noise in the data.

In order to evaluate the reliability of the sample discrimination, confidence ellipses were applied to compare differences among the three Napping configurations. The classical Napping had the largest ellipses; *Napping_Method* showed smaller ellipses and *Napping_Product* showed the smallest ellipses. Moreover, in the classical Napping plot most ellipses overlapped with each other, indicating poor discrimination between the samples. However, when observing the positioning of the wine samples in this plot (Fig. 2a), it can be seen that the samples with the same added flavour compound were close to each other, especially for the isopentyl acetate group which had the smallest distance between two samples with different concentrations. The *Napping_Method* plot (Fig. 2b) showed a clearer grouping among the wine samples, where there was no overlap between 2-phenylethanol group and the other two groups (benzaldehyde and isopentyl acetate group). The *Napping_Product* plot (Fig. 2c) displayed three clearly separate groups, among which there were no overlaps. Overall, the positioning of samples demonstrated that conducting training strategies provided more reliable Napping results. However, in all the Napping plots the two samples with the same added compound but in different concentrations were not separated from each other (e.g. Iso High and Iso Low). The base wine, moreover, could not be separated from the 2-phenylethanol group in all Napping approaches, which was in line with the scaling data showing that the 2-phenylethanol was the most difficult to be distinguished from the base wine.

As mentioned previously, the positioning of blind duplicates was considered as a way to check the panel performance and examine the accuracy of the methods. A higher level of accuracy was observed in the two Napping approaches with training as expected. The blind duplicates of 2-phenylethanol and benzaldehyde were located closer to each other in the *Napping_Method* and *Napping_Product* plots compared to the classical Napping, indicating that both types of training had a positive effect on the accuracy of the panel description.

3.2.2. Vocabulary comparison

Attributes were collected from subsequent attributes generation steps in all three Napping approaches. The attributes and their frequencies of mention are shown in Table 4. Attributes mentioned less than six times were not shown to improve the interpretability

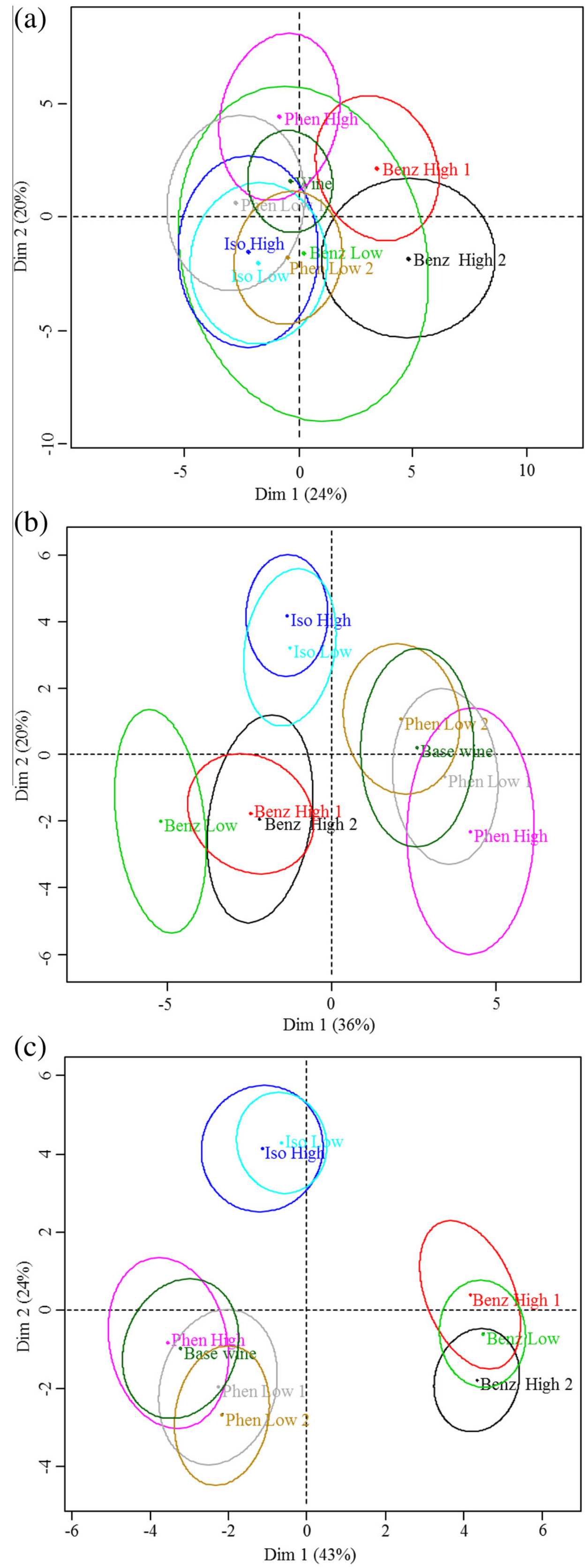

Fig. 2. MFA product plots with confidence ellipses obtained from Napping. (a) Classical Napping. (b) Napping with training on method. (c) Napping with training on product. The samples with additions of benzaldehyde, isopentyl acetate and 2-phenylethanol are abbreviated to Benz, Iso and Phen, respectively. Low and High mean that the flavour compounds were added in 'low' and 'high' concentration, respectively.

as suggested in the literatures (Perrin & Pagès, 2009; Reinbach, Giacalone, Ribeiro, Bredie, & Frøst, 2014). Among these attributes almond, fruity, floral and banana were commonly mentioned by all panels, which may indicate that the core set of sensory attributes matched the usual descriptions of the compounds added to the model wine. It was observed that the number of total attributes reduced and the frequencies of core attributes were larger when conducting training strategies: a total of 59, 58 and 42 attributes were obtained for the classical Napping, *Napping_Method* and *Napping_Product* approaches, respectively. The most frequently mentioned attributes were almond (61), banana (43) and floral (32) in the *Napping_Product*, almond (37) and mild (30) in the *Napping_Method*, and almond (34) in the classical Napping. These results suggest that especially the product training was effective in producing a more semantically clear and coherent data structure.

3.3. Flash Profile

3.3.1. Comparison of configurations obtained with the two FP approaches

Fig. 3 displays the MFA results performed on the data from both FP approaches (averaged over the two replicates). Unlike Napping, two FP approaches produced the same total explained variance of the first two dimensions (47%) of their respective MFA models. The modified FP had a slightly more bi-dimensional structure, with a lower variance for the first dimension (26%) and a higher variance for the second dimension (21%) compared to the classical FP (30% and 17%, respectively). There were no clear groups for samples with the same added flavour compounds as Napping with training (Fig. 2b and c), however, the intensity differences between samples were more visible. Accordingly, both FP plots showed that the wine samples with low concentrations of added flavour compounds were located around the centre, which was close to the base wine, while the samples with high concentrations were located further from the centre. Moreover, there were no overlaps among the samples with high concentrations of added flavour compounds in both plots, indicating these samples were significantly different from each other in both FP approaches. However, the base wine appeared more clearly separated from the other samples in the modified FP (Fig. 3b) compared to the classical FP plot (Fig. 3a). It should be noted that both plots showed that the samples with added compounds in low concentrations were not significantly discriminated from each other.

The positioning of the two pairs of blind duplicates was also examined. The 2-phenylethanol low blind duplicate got closer to each other in the modified FP plot, however, the positioning of the duplicated benzaldehyde did not show any improvement compared to the classical FP.

3.3.2. Vocabulary comparison

The attributes generated in the two FP approaches are listed in Table 5. The total number of attributes excluding repeated citation for the classical FP was 34 with each assessor having between 6 and 13 attributes; the number of attributes for the modified FP was, as expected, lower and amounted to 22 with each assessor having between 5 and 9 attributes. Limiting attributes that the assessors could use for ranking was expected to contribute to focusing on the major characteristics of products with less interference of inter-individual variation.

3.4. Repeatability between sessions for both Napping and FP

As each method was carried out in duplicate in this study, it was possible to investigate the agreement between replicates, and use the RV coefficient to infer the method repeatability. The lowest RV coefficient between replicates 0.59 ($p = 0.31$) was obtained for the classical Napping, although overall similarities and differences among samples in the plot were constant over repeated sessions. An improvement of repeatability was observed when introducing a training step in Napping: the RV coefficients were 0.68 ($p = 0.012$) for the *Napping_Method* and 0.85 ($p = 0.001$) for the *Napping_Product*, respectively. As observed for the discriminability, also with regards to repeatability adding a small training step improved the Napping performance and, again, the product training was superior to the method training. The RV coefficients between replicates for the classical FP and modified FP were 0.84 ($p = 0.001$) and 0.87 ($p = 0.001$), respectively, indicating a good level of repeatability of both protocols.

4. Discussion

4.1. The effect of training strategies on Napping

One of the aims of this study was to evaluate the effect of adding a training step to a Napping task. Two types of training were considered. The first was a training on the method itself, as proposed by Hopfer and Heymann (2013), aimed at ensuring that all assessors have a good understanding on the method. The second was a training step on the products, inspired by Perrin et al. (2008), who suggested that the Napping methodology could be modified by proposing a list of attributes to associate to the products in advance.

Table 4
Attributes mentioned more than five times by assessors in three Napping approaches.

Classical Napping		*Napping_Method*		*Napping_Product*	
Attributes	Frequency of mention	Attributes	Frequency of mention	Attributes	Frequency of mention
Almond	34	Almond	37	Almond	61
Fruity	21	Mild	30	Banana	43
Floral	20	Floral	23	Floral	32
Honey	20	Fruity	22	Neutral	23
Banana	16	Strong	21	Strong	20
Pear	15	Peach	13	Fruity	18
Citrus	15	Apple	13	Mild	10
Elderflower	14	Chemical	10	Fatty	8
Mild	14	Citrus	8	Apple	7
Dried apricot	12	Banana	8	Pear	6
Rose	11	Spicy	7		
Neutral	9	Pear	6		
Berry	8	Neutral	6		
Fresh	8	Cinnamon	6		
Vinegar	6				

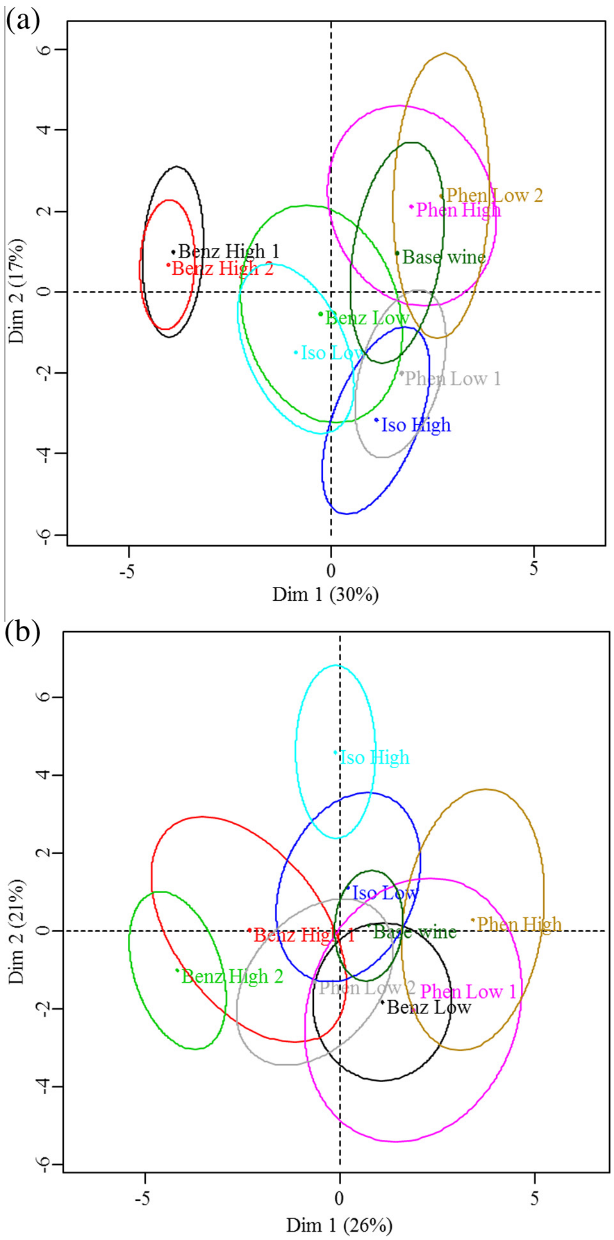

Fig. 3. MFA product plots with confidence ellipses obtained from Flash Profile. (a) Classical Flash Profile, and (b) modified Flash Profile. The samples with additions of benzaldehyde, isopentyl acetate and 2-phenylethanol are abbreviated to Benz, Iso and Phen, respectively. Low and High mean that the flavour compounds were added in 'low' and 'high' concentration, respectively.

When comparing the results of both panels who received training with the classical Napping, results clearly showed that training had a positive effect on the discriminability, with the product training being more effective than the method training. The superiority of product training seems well in line with previous findings that familiarity with the product leads to a better discrimination performance in Napping tasks (Giacalone, Machado, & Frøst, 2013; Torri et al., 2013), suggesting that Napping is a valuable method particularly when assessors share a high level of knowledge and experience about the product. However, it is worth noting that when complex products with more different characteristics are used, and these differences are not exactly known beforehand, more products covering a larger span in sensory differences will need to be present in the familiarisation step.

The training also resulted in a higher repeatability. A higher RV coefficient between two repeated sessions was observed for both *Napping_Method* and *Napping_Product*. This is also consistent with earlier findings that the repeatability of rapid methods is positively affected by training (Ares & Varela, 2014, chap. 14). Considering that this study employed a model food system, in real products applications we still suggest to perform replicated sessions when running a Napping task, even with training protocols, to ensure repeatability and validity of the results. As recommended by Mielby et al. (2014), using replicates for fast sensory methods is advisable also because one could not make sure that a higher number of assessors in one session would give the same result as replicates.

For all these reasons, we suggest the use of product training when conducting Napping with trained panels as an alternative to descriptive analysis. It should be noted that product training may not be desirable in all applications. For example, when consumers are used for the evaluation it might not be of interest to direct their attention to pre/determined attributes, but rather to let the panel decide what the most important sensory characteristics are.

4.2. The effect of modifications on FP

A second aim of this work was to evaluate a modified FP protocol compared to the classical version. Our first modification on the FP was the addition of a Napping as the attribute generation step prior to the FP task. This was inspired by Perrin and Pagès (2009), who proposed that attributes collected in the context of a holistic sensory task such as Napping could be more consistent than a totally free description. Moreover, the positioning task could be a familiarising step for the assessors. A second modification on the FP was limiting the number of attributes that assessors could use for ranking. It was hypothesised that the modification could help the assessors focus on the core differences between samples, decrease the noise level in the sample ranking and thus improve the semantic interpretability of the results. The choice of setting the limit at 10 attributes was because this number allowed an adequate number of sensory dimensions while limiting the assessor to choosing the most relevant to the products considered. In future studies, the exact number could vary depending on the sample complexity, the objectives of the study, etc.

Unlike the Napping results, the modifications to the FP approach did not produce a significant improvement compared to the original protocol, as both the classical and modified FP provided similar results. However, we expect that this result is partly due to the samples themselves, since in this work we used model wines with limited sensory differences. The comparative advantage of adding a familiarisation step and limiting the number of attributes might be larger if more complex products were used, so in real applications more pronounced improvements in sample discrimination through the proposed modifications may be expected.

4.3. Comparison of Napping and Flash Profile

When comparing the two types of approaches, it was observed in present study that Napping was mainly able to catch character differences among wine samples with different flavour compounds added, while the FP approaches could also catch intensity differences between wine samples with the same added flavours but in different concentrations. This result is consistent with the different nature of the respective tasks. Napping is a holistic methodology where the assessors focus on global differences and similarities among all samples, which enables them to identify the most salient sensory characteristics of the samples. On the other hand, FP, as an

Table 5
Attributes used by assessors in two Flash Profile approaches.

Classical Flash Profile				Modified Flash Profile			
Attributes	Frequency of mention	Attributes	Frequency of mention	Attributes	Frequency of mention	Attributes	Frequency of mention
Almond	7	Vanilla	1	Almond	7	Rose	1
Floral	5	Apricot	1	Banana	5	Nutty	1
Elderflower	5	Cherry	1	Citrus	5	Neutral	1
Banana	4	Fresh	1	Pear	5	Vinegar	1
Fruity	3	Dust	1	Apricot	4	Corky	1
Muscat	3	Grapefruit	1	Fruity	4	Ash	1
Woody	3	Grape	1	Honey	4		
Plum	3	Grass	1	Berry	3		
Honey	2	Juniper	1	Cheesy	3		
Apple	2	Lemon	1	Elderflower	3		
Citrus	2	Mild	1	Floral	3		
Green apple	2	Mushroom	1	Strawberry	3		
Lime	2	Pear	1	Acetone	2		
Peach	2	Peppermint	1	Woody	2		
Neutral	2	Pine	1	Spicy	2		
Raisin	2	Spicy	1	Raisin	1		
Rose	2	Sulphur	1				

attribute-based methodology, focuses the assessors' attention on specific attributes thus provides more detailed information on samples (Valentin, Chollet, Lelièvre, & Abdi, 2012; Varela & Ares, 2012). Accordingly, Fig. 3 shows that, in both versions of the FP, the samples with the high flavour intensity were located further away from the centre of the plot, whereas the base wine and the samples with low intensity were located close to the origo. This was unlike the results from the Napping panels (Fig. 2), in which the samples appeared to be mainly grouped according to the dominant characters, without discrimination between intensities. This is also generally consistent with previous studies comparing holistic methodologies and attribute-based methodologies in different products categories. For example, in a study on hot beverages Moussaoui and Varela (2010) found that FP and RGM gained benefit of providing a quantitative product mapping compared to Sorting and Projective Mapping, although all methods could be applied as a means of generating vocabulary.

If we look at the size of sensory ellipses produced by the two kinds of methodologies (Figs. 2 and 3), the classical Napping and the two FP approaches showed the largest degree of overlap, and thus lower discriminability. While it is fair to say that the classical Napping performed worse than the Napping panels with training, we are more cautious in using confidence ellipses to compare Napping and FP. There is an important caveat here: if we consider the compositions of the wine samples (Table 1), it is clear that specific attributes could actually be perceived in only two or three out of nine samples. However, the FP protocol requires assessors to rank all samples for each attribute and indeed they mostly ideally ranked the first two samples, but were essentially guessing afterwards. This may have inflated the noise level in the data and therefore resulted in large confidence ellipses. Besides, the relatively small size of the panels might also contribute to large confidence ellipses. It is therefore still necessary to study the influence of the number of assessors on the performance of different rapid methodologies.

4.4. Positioning of blind duplicates within session for both Napping and FP

Hopfer and Heymann (2013), Nestrud and Lawless (2010), Moussaoui and Varela (2010), and Veinand et al. (2011) evaluated the accuracy of rapid methodologies by comparing the position of a blind repeated sample on the product space. These authors reported that repeated samples were located close to each other in the sample space and concluded that Napping and FP were reliable. However, it is important to take into account that when working with very similar samples, blind repeated sample could be placed apart from each other in the product space, which would not necessarily imply lack of accuracy (Ares & Varela, 2014, chap. 14). In our work, it was observed that the accuracy of Napping improved markedly when conducting training, as both *Napping_Method* and *Napping_Product* panels placed the blind duplicates closer to each other compared to the panel who did not receive training. For FP, both protocols showed a relatively good accuracy, but no large differences were observed between the classical and the modified version.

5. Conclusions

This work explored whether small adjustments to rapid sensory methodologies could improve the results. Napping and Flash Profile were chosen as methods to test discriminability, repeatability and accuracy. Model white wines were developed with known differences on the principal sensory properties and intensities, varying from weak to moderate. Regarding Napping, the results showed that conducting training on either the method or product provided more robust results, compared to the classical protocol, where the assessors did not received any extra training. Moreover, the product training was generally more effective than the training on method prior to Napping. The classical Flash Profile method did not differ, in terms of sample discrimination and accuracy, from the modified Flash Profile method, which included a Napping with subsequent attributes generation as the word generation step and a limitation on the number of attributes in the product ranking. Comparing Napping and Flash Profile, this study showed that Napping mainly provided qualitative information on the dominant flavour character, while the Flash Profile provided both qualitative and quantitative information, in line with the known intensity differences of the flavours in the wines. This may be expected due to the different rationales of the two methodologies, where one (Napping) utilises sensory similarities and differences as discrimination criteria, while the other (FP) is based on the evaluation of verbalized sensory attributes and also includes intensity ranking for sample discrimination. Nevertheless, the overlap in confidence ellipses between low and high intensity samples in the FP plots do not completely support this interpretation, suggesting that this aspect should be given further consideration in future work.

Overall, this study showed that minor modifications to existing protocols could significantly change the quality of the results obtained by rapid sensory methods. In this paper, this was mostly the case for Napping. More generally, this work also suggests that there is a need for further studies to develop, validate, and improve rapid sensory methodologies, to reach a compromise between the laborious and slow conventional descriptive analysis, and the very fast but less robust alternative methodologies.

Acknowledgments

This work was financially supported by the University of Copenhagen – Denmark and Chinese Scholarship Council. Author R. Di Monaco acknowledges support by Campania Region fund within the program "POR CAMPANIA FSE 2007/2013" – project CARINA. The authors would like to thank Mehtap Tekin for her assistance of part of the lab work.

References

Albert, A., Varela, P., Salvador, A., Hough, G., & Fiszman, S. (2011). Overcoming the issues in the sensory description of hot served food with a complex texture. Application of QDA®, flash profiling and projective mapping using panels with different degrees of training. *Food Quality and Preference, 22*, 463–473.
Ares, G., & Varela, P. (2014). Comparison of novel methodologies for sensory characterization. In G. Ares & P. Varela (Eds.), *Novel techniques in sensory characterization and consumer profiling*. Bosa Roca, Hoboken: Taylor and Francis.
Barcenas, P., Elortondo, F. J. P., & 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.
Byrne, D. V., Bredie, W. L. P., & Martens, M. (1999). Development of a sensory vocabulary for warmed-over flavour. Part II: In chicken meat. *Journal of Sensory Studies, 14*, 67–78.
Byrne, D. V., O'Sullivan, M. G., Dijksterhuis, G. B., Bredie, W. L. P., & Martens, M. (2001). Sensory panel consistency during development of a vocabulary for warmed-over flavour. *Food Quality and Preference, 12*, 171–187.
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*, 826–834.
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., Brockhoff, P. B., Meinert, L., Aaslyng, M. D., & Bredie, W. L. P. (2012). Rapid descriptive sensory methods – Comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference, 26*, 267–277.
Delarue, J., & Sieffermann, J.-M. (2004). Sensory mapping using Flash profile. Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. *Food Quality and Preference, 15*, 383–392.
Escoufier, Y. (1973). Le traitement des variables vectorielles. *Biometrics (International Biometric Society), 29*, 751–760.
Giacalone, D., Machado, L., & Frøst, M. B. (2013). Consumer-based sensory profiling: Application of Partial Napping® for Sensory Characterization of Specialty Beers by Novices and Experts. *Journal of Food Products Marketing, 19*, 201–218.
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.
ISO-3591 (1977). *Sensory analysis – Apparatus – Wine-tasting glass*. Geneva, Switzerland: International Standardization Office.
Josse, J., Pagès, J., & Husson, F. (2008). Testing the significance of the RV coefficient. *Computational Statistics & Data Analysis, 53*, 82–91.
Kennedy, J. (2010). Evaluation of replicated projective mapping of granola bars. *Journal of Sensory Studies, 25*, 672–684.
Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling: Newbury Park*. CA: Sage.
Lê, S., & Husson, F. (2008). SensoMineR: A package for sensory data analysis. *Journal of Sensory Studies, 23*, 14–25.
Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software, 25*, 1–18.
Liu, J., Toldam-Andersen, T. B., Petersen, M. A., Zhang, S., Arneborg, N., & Bredie, W. L. P. (2015). Instrumental and sensory characterisation of Solaris white wines in Denmark. *Food Chemistry, 166*, 133–142.
Mielby, L. H., Hopfer, H., Jensen, S., Thybo, A. K., & Heymann, H. (2014). Comparison of descriptive analysis, projective mapping and sorting performed on pictures of fruit and vegetable mixes. *Food Quality and Preference, 35*, 86–94.
Moussaoui, K. A., & Varela, P. (2010). Exploring consumer product profiling techniques and their linkage to a quantitative descriptive analysis. *Food Quality and Preference, 21*, 1088–1099.
Nestrud, M. A., & Lawless, H. T. (2008). Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. *Food Quality and Preference, 19*, 431–438.
Nestrud, M. A., & Lawless, H. T. (2010). Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies, 25*, 390–405.
Pagès, J. (2003). Direct collection of sensory distances application to the evaluation of ten white wines of the Loire Valley. *Sciences des Aliments, 23*, 679–688.
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.
Perrin, L., & Pagès, J. (2009). Construction of a product space from the ultra-flash profiling method: Application to 10 red wines from the Loire Valley. *Journal of Sensory Studies, 24*, 372–395.
Perrin, L., Symoneaux, R., Maître, I., Asselin, C., Jourjon, F., & Pagès, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: Case of ten wines from Loire Valley. *Food Quality and Preference, 19*, 1–11.
Price, E. J., Linforth, R. S., Dodd, C. E., Phillips, C. A., Hewson, L., Hort, J., et al. (2014). Study of the influence of yeast inoculum concentration (*Yarrowia lipolytica* and *Kluyveromyces lactis*) on blue cheese aroma development using microbiological models. *Food Chemistry, 145*, 464–472.
R Development Core Team (2010). *R: A language and environment for statistical computing. 3-900051-07-0*. Vienna, Austria: R Foundation for Statistical Computing.
Reinbach, H. C., Giacalone, D., Ribeiro, L. M., Bredie, W. L. P., & Frøst, M. B. (2014). Comparison of three sensory profiling methods based on consumer perception: CATA, CATA with intensity and Napping®. *Food Quality and Preference, 32*, 160–166.
Risvik, E., McEwan, J. A., Colwill, J. S., Rogers, R., & Lyon, D. H. (1994). Projective mapping: A tool for sensory analysis and consumer research. *Food Quality and Preference, 5*, 263–269.
Risvik, E., McEwan, J. A., & Rødbotten, M. (1997). Evaluation of sensory profiling and projective mapping data. *Food Quality and Preference, 8*, 63–71.
Sieffermann, J. M. (2000). Le profil Flash: Un outil rapide et innovant d'é valuation sensorielle descriptive. In *Paper presented at the L'innovation: De l'idé eau succés – Douziè mes rencontres AGORAL*. Paris, France: Lavoisier, Tec & Doc.
Torri, L., Dinnella, C., Recchia, A., Næs, T., Tuorila, H., & Monteleone, E. (2013). Projective mapping for interpreting wine aroma as perceived by naïve and experienced assessors. *Food Quality and Preference, 29*, 6–15.
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 & Technology, 47*, 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.
Veinand, B., Godefroy, C., Adam, C., & Delarue, J. (2011). Highlight of important product characteristics for consumers. Comparison of three sensory descriptive methods performed by consumers. *Food Quality and Preference, 22*, 474–485.
Williams, A. A., & Langron, S. P. (1984). The use of free-choice profiling for the evaluation of commercial ports. *Journal of the Science of Food and Agriculture, 35*, 558–568.