Food Quality and Preference 32 (2014) 160–166

Contents lists available at SciVerse ScienceDirect

Food Quality and Preference

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

# Comparison of three sensory profiling methods based on consumer perception: CATA, CATA with intensity and Napping®

Helene C. Reinbach [1], Davide Giacalone [1], Leticia Machado Ribeiro, Wender L.P. Bredie, Michael Bom Frøst *

*Department of Food Science, Faculty of Science, University of Copenhagen, Rolighedsvej 30, 1958 Frederiksberg C, Denmark*

A R T I C L E   I N F O

*Article history:*
Received 26 October 2012
Received in revised form 22 February 2013
Accepted 25 February 2013
Available online 7 March 2013

*Keywords:*
Fast sensory methods
Napping
Check-all-that-apply
Descriptive profiling
Multivariate statistics
Consumer perception

A B S T R A C T

The present study compares three profiling methods based on consumer perceptions in their ability to discriminate and describe eight beers. Consumers ($n = 135$) evaluated eight different beers using Check-All-That-Apply (CATA) methodology in two variations, with ($n = 63$) and without ($n = 73$) rating the intensity of the checked descriptors. With CATA, consumers rated 38 descriptors grouped in seven overall categories (berries, floral, hoppy, nutty, roasted, spicy/herbal and woody). Additionally 40 of the consumers evaluated the same samples by partial Napping® followed by Ultra Flash Profiling (UFP). ANOVA- and Discriminant Partial Least Square Regression (A-PLSR, D-PLSR) were used to evaluate the discriminative ability of the methods and descriptors. A-PLSR results showed that all samples were perceived as different in all three methods, whereas D-PLSR showed that all three methods had similar numbers of discriminating descriptors. For the two CATA variants, 29 and 24 descriptors for without and with rating intensity were significant, for Napping/UFP the number was 26. Multiple Factor Analysis was used to derive an overall product map and to compare it to product configurations from individual methods. Both qualitative and quantitative analysis (comparison of $R_V$ coefficients of the MFA configurations) revealed a very high agreement of the three methods in terms of perceived product differences. $R_V$ coefficients were used to compare sample configurations obtained in the three descriptive methods. For all comparisons the $R_V$ coefficients varied between 0.90 and 0.97, indicating a very high similarity between all three methods. These results show that the precision and reproducibility of sensory information obtained by consumers by CATA is comparable to that of Napping. The choice of methodology for consumer descriptive methods should then be based on whether it is desired to have consumers articulate their own perception of descriptors, or if it sufficient to present them to an existing vocabulary. Napping is slower and more laborious, and better for explorative studies with smaller number of consumers whereas, CATA is faster, less labor-intensive and thus more suitable for larger groups of consumers.



## 1. Introduction

Descriptive sensory profiling is important for the food industry as it can guide product development and reformulation of products as well as identify key sensory drivers essential for consumer acceptance and marketing of products. Conventional descriptive profiling is performed with a trained panel to obtain an objective description of the food products investigated (Lawless & Heymann, 2010). The need for less time-consuming and economical descriptive methods in the food industry has supported the development and use of more dynamic and fast descriptive sensory profiling methods assessed by panelists, food experts and consumers (Ares, Deliza, Barreiro, Gimenez, & Gambaro, 2010; Dehlholm, Brockhoff, Meinert, Aaslyng, & Bredie, 2012; Giacalone, Machado Ribeiro, & Frøst, forthcoming; Nestrud & Lawless, 2010). The fast methods include projective mapping (Risvik, McEwan, & Rødbotten, 1997) and Napping® (Pages, 2003, 2005), Flash Profiling (Dairou & Sieffermann, 2002) based on Free-Choice Profiling (Williams & Langron, 1984) and different sorting techniques such as free (Lawless, Sheng, & Knoops, 1995) single (Rosenberg & Kim, 1975) and multiple sorting (Dehlholm et al., 2012). Napping® is a method in which food samples are projected on a two-dimensional space based on similarities, and is often combined with Ultra Flash Profiling (Perrin & Pages, 2009) to add a semantic description to the product differences. Napping can performed as a "global" Napping, including all sensory aspects, or as "partial" Napping focusing on specific sensory modalities (e.g. appearance, taste or mouthfeel) (Dehlholm et al., 2012; Pagès, 2005).

Other consumer-friendly methods, such as just-about right scales (JAR), attribute liking, emotional questionnaires and check-all-that-apply (CATA) are increasingly used to capture consumer perception of food products. In particular the CATA method, in

* Corresponding author. Tel.: +45 35 33 32 07; fax: +45 35 33 3509.
*E-mail address:* mbf@life.ku.dk (M.B. Frøst).
[1] These authors contributed equally to this work.

0950-3293/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodqual.2013.02.004

which a product is described by selecting appropriate words from a given list, is a simple and valid approach to gather information about sensory and non-sensory perception, and is believed to have smaller effect on liking and consumer perception of the product than similar methods (e.g. JAR) (Adams, Williams, Lancaster, & Foley, 2007; Ares et al., 2010; Giacalone, Bredie, & Frøst, 2013; Lado, Vicente, Manzoni, & Ares, 2010). Consumer-elicited CATA profiles have shown good agreement with traditional panel-developed sensory profiles (Dooley, Lee, & Meullenet, 2010; Ares et al., 2010), suggesting that CATA could be a valuable alternative to understand perception of product sensory attributes.

The various methodologies to capture consumer perceptions are generally easier to perform and less time-consuming than traditional descriptive analysis with a trained sensory panel. Some methods are reductionist and based on a predefined list of descriptor (e.g. CATA), while other methods are more holistic and explorative (e.g. Napping). One of the suggested drawbacks of CATA is that this method produces relatively impoverished dichotomized data (1/0), which allegedly would mask relative differences between specific attributes. Including intensity scaling of attributes in the CATA method may therefore improve the accuracy of descriptive profiling and lead to a better product differentiation. This hypothesis could be tested by comparing CATA with CATA combined with intensity scaling. Data on consumer ratings of intensity generally show large variability and thus it is not clear if the scaling element would actually improve the CATA descriptions made by the consumer. Additionally, it would be of interest to compare how reductionist methods, with and without scale elements (CATA and CATA with intensity ratings), would fare compared to a more holistic and explorative one, such as Napping.

The aim of the present study was to compare the effectiveness of three profiling methods, CATA, Napping and a novel method combining CATA with intensity scaling in studying consumer perception of a sample of eight beers. Three comparative criteria were considered in this study:

(1) *Discriminative ability*: i.e. the method's ability to successfully discriminate between the samples;
(2) *Descriptive ability*: the degree to which the three profiling methods would agree on the sensory characterization;
(3) *Configurational congruence*: the degree to which the sample spaces obtained by the different methods would be closely related to one another.

## 2. Materials and methods

### 2.1. Consumers

One hundred and thirty-five consumers between 18 and 65 years were recruited in and around of University of Copenhagen (UCPH), through advertisement on websites, social networks, beer magazines and flyers. Approximately half of the consumers ($n = 73$, 46 males and 27 females) described the flavor of the beers using a CATA questionnaire. The other half ($n = 62$, 46 males, 16 females) completed a modified version of the CATA questionnaire where we introduced the possibility of scaling the intensity of checked attributes. Additionally, some of the consumers ($n = 40$, 23 males, 17 females) returned after approximately 10 days for a second session to perform a partial Napping focusing on the smell and taste attributes of the eight beers.

After the testing, consumers received a token incentive for their participation (a bottle of craft beer, value ≈ 6 €).

### 2.2. Samples

Eight beers were chosen for the study (Table 1), five that represented the flavor diversity of the Danish beer marked (e.g. fruity, floral, woody, nutty or spicy), two beers were developed for the study to represent novel ingredients (sea buckthorn and pine) and finally a standard pilsner was included to represent the most consumed beer type in Denmark.

40 ml of beer was served at approximately 10 °C in 24 cl beer glass covered with watch glasses and coded with three-digit random numbers. Serving orders were randomized to balance first order and carry-over effects (MacFie, Bratchell, Greenhoff, & Vallis, 1989).

### 2.3. CATA variants

Sensory perception of the eight beers was evaluated by respectively CATA and CATA combined with a 15-point intensity scale. On the CATA ballot seven overall flavor categories were presented (Table 2). For each flavor category consumers were asked to check *yes*, if the flavor was present, and *no* if the flavor was not present. This formulation[2] differs from the classical "check-all-that-apply", and was adopted in order to enhance the likelihood that consumers actually read through the whole list, reducing the behavior known as *satisficing* (Krosnick, 1991; Rasinski, Mingay, & Bradburn, 1994). Briefly, satisficing is a theory in behavioral decision making maintaining that when most people examine alternative sequentially, they tend to choose the first alternative that seem reasonable, as opposed to the optimal situation in which they would evaluate all alternatives comprehensively before taking a decision (Simon, 1955).

Further, some overall flavor descriptors were supplemented with sub-descriptors to enable consumers to specify the exact flavor they perceived (Table 2). The list of flavor attributes was developed with inspiration from the "Danish beer language" (Det Danske Ølakademi [Eng. *The Danish Beer Academy*], 2006), and the ballot was pre-tested informally to assess that the appropriateness of the attribute list. On the CATA ballot with intensity scaling, the seven overall flavor attributes were presented with the *yes*/*no* checkboxes, the flavor sub-descriptors and one horizontally oriented 15-point intensity scale per flavor category anchored with 'very weak' and 'very strong' in the ends to enable consumers to rate the intensity of the appropriate beer flavors. The choice of including only flavor terms, which differs from earlier CATA applications where often more holistic terms (e.g. emotions, usage attributes, conceptual attributes, etc.) are included, was motivated by our aim to restrict the focus on the descriptive profiling applicability of CATA.

### 2.4. Partial Napping

Napping was performed as a partial Napping focusing on the smell and taste of the eight beers. Each consumer was provided with a 60 × 40 cm blank paper (the Napping sheet), a pen, post-its, a tray with eight beer samples and a spittoon. The sample order on the individual trays was randomized to counter-act first order carry-over effect, even though the Napping methodology allows and requires subjects to go back and forth between samples. Consumers were instructed to evaluate the beer samples according to similarities or dissimilarities in smell and taste attributes by placing similar samples close to each other and more dissimilar samples further apart on the Napping sheet. After they had reached a final configuration, consumers noted down appropriate descriptors for the smells and tastes of the beers on the post-its, which were moved around the Napping sheet, when needed. This procedure is known as Ultra-Flash profiling and is commonly used to add a descriptive dimension to a Napping task (Perrin et al., 2008). When

[2] A very similar formulation has been recently tested by Ennis and Ennis (2011), who coined their approach "applicability scores". Although unaware of this contribution at the time of designing this experiment, it is interesting to notice that we came to very similar conclusions regarding the need to account for unchecked items in CATA questionnaires.

**Table 1**
Description of the eight beers.

| Beer name | Main sensory characteristics | Main flavor ingredient | Beer style | Brewery |
|---|---|---|---|---|
| Sea buckthorn beer | Berries | Sea buckthorn | Flavored pilsner | Indslev/Univ. of Copenhagen (UCPH) |
| Fynsk forår | Floral | Elderflower | Pale Ale | Ørbæk |
| Bøgebryg | Woody | Beech tree extract | Amber Ale | Skovlyst |
| Pine beer | Woody | Pine (Pin-Thyrol) | Flavored pilsner | Indslev/UCPH |
| Valnød Hertug | Nutty | Walnut | Dark Ale | Rise |
| Stjernebryg | Spicy | Anise | Strong Ale | Herslev |
| Enebær Stout | Spicy | Juniper berries | Stout | Grauballe |
| Thy Pilsner | Neutral | Hops | Pilsner | Thy |

**Table 2**
Overview of descriptors for the three methods and estimated jack-knife significance from D-PLSR.

| Common descriptors | CATA ($n = 73$) | CATA Int. ($n = 62$) | Napping ($n = 40$) | Other Napping descriptors (by # of occurrences) |
|---|---|---|---|---|
| BERRIES | $166^{**}$ | $140^{**}$ | $7^{n.s.}$ | $\text{Sweet}^{***}$ (96) |
| Blueberry | $16^{*}$ | $9^{*}$ | | $\text{Bitter}^{**}$ (74) |
| Cranberry | n.s. | $31^{*}$ | | $\text{Sour}^{***}$ (48) |
| Sea buckthorn | n.s. | n.s. | | $\text{Citrus}^{***}$ (24) |
| Rose hip | $58^{**}$ | n.s. | | $\text{Fresh}^{**}$ (22) |
| Other berries | n.s. | $21^{**}$ | | $\text{Fruity}^{n.s.}$ (20) |
| Floral | $242^{***}$ | $215^{***}$ | | $\text{Strong}^{***}$ (20) |
| Elderflower | $93^{***}$ | $77^{***}$ | $22^{***}$ | $\text{Light}^{*}$ (19) |
| Chamomile | $50^{**}$ | $38^{**}$ | | $\text{Liquorice}^{**}$ (19) |
| Lavender | $61^{**}$ | n.s. | | $\text{Yeasty}^{**}$ (20) |
| Rose | $36^{**}$ | n.s. | | $\text{Full-bodied}^{*}$ (18) |
| Other floral | $27^{**}$ | n.s. | | $\text{Alcoholic}^{**}$ (15) |
| HOPPY | $345^{***}$ | $389^{***}$ | $45^{***}$ | $\text{Pilsner}^{**}$ (15) |
| NUTTY | $200^{***}$ | $149^{***}$ | $18^{*}$ | $\text{Summer}^{*}$ (14) |
| Hazelnut | $72^{***}$ | $47^{*}$ | | $\text{Thin}^{*}$ (13) |
| Almond | $47^{***}$ | $33^{*}$ | | $\text{Burnt}^{*}$ (12) |
| Walnut | $71^{***}$ | $48^{***}$ | $9^{n.s.}$ | $\text{Grainy}^{n.s.}$ (12) |
| Other nutty | n.s. | | | $\text{Watery}^{n.s.}$ (11) |
| ROASTED | $271^{***}$ | $243^{***}$ | $10^{*}$ | $\text{Heavy}^{n.s.}$ (9) |
| Roasted bread | $60^{*}$ | n.s. | | $\text{Soapy}^{*}$ (8) |
| Caramel | $133^{***}$ | $111^{***}$ | $33^{***}$ | $\text{Neutral}^{n.s.}$ (8) |
| Coffee | $64^{***}$ | $64^{***}$ | $8^{*}$ | $\text{Spring}^{n.s.}$ (7) |
| Chocolate | $36^{***}$ | $47^{***}$ | $8^{*}$ | $\text{Regular}^{*}$ (6) |
| Other roasted | n.s. | $17^{*}$ | | $\text{Autumn}^{n.s.}$ (6) |
| SPICY/HERBAL | $295^{***}$ | $274^{***}$ | $19^{*}$ | $\text{Apple}^{n.s.}$ (5) |
| Juniper berries | $52^{*}$ | $50^{*}$ | | $\text{Low bitterness}^{n.s.}$ (5) |
| Bog myrtle | n.s. | $70^{*}$ | | |
| Anise | $56^{**}$ | $64^{***}$ | 8 | |
| Rosemary | $47^{***}$ | $35^{***}$ | | |
| Cloves | $44^{*}$ | n.s. | | |
| Laurel | n.s. | n.s. | | |
| Other spicy/herbal | n.s. | n.s. | | |
| WOODY | $232^{**}$ | $205^{***}$ | $12^{n.s.}$ | |
| Piney | $70^{**}$ | $71^{***}$ | $19^{**}$ | |
| Birch | $60^{*}$ | n.s. | | |
| Beech | $33^{*}$ | n.s. | | |
| Maple | $36^{*}$ | n.s. | | |
| Other woody | n.s. | n.s. | | |
| Significant descriptors | 29 | 24 | 26 | |
| Optimal PC | 4 | 4 | 5 | |
| Validated *Y* variance | 16% | 11% | 22% | |

*p* Values for $\beta$ coeff. Significance levels, *n.s.* non-significant.
* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.

all the samples had been placed on the paper, the assessors substituted the post-its with an *X* and noted the sample codes and the beer characteristics next to the *X*. Each beer sample was tasted in the given order and swallowed at least once to get the full perception of the beers. Re-tastings and spitting out the beer were allowed. Water and crackers were used as palate cleansers.

### 2.5. Data analysis

The analysis conducted was divided into two parts: an analysis of the descriptive outputs of the three methods (using unfolded data matrices and Partial Least Squares Regression), and an analysis of the sample spaces (using crosstab matrices and Multiple Factor Analysis).

The descriptive profiles of the beers obtained from CATA, CATA with intensity and Napping were analyzed for product differences by ANOVA-Partial least square regression (A-PLSR; Martens & Martens, 2001) using the Unscrambler software (version 10.1, CAMO, Oslo, Norway). The matrices thus consisted of the number of beers (8) times the number of consumers in each group. For the A-PLSR analysis, the *X*-matrix consisted of the eight experimental beers ($X = 1/0$ design variables) while the *Y*-matrix consisted of the beer flavor descriptors for CATA ($Y = 1/0$) and CATA with intensity ($Y = \text{intensity}/0$). The *Y*-matrix for the Napping included the taste and flavor descriptors of the beers ($Y = 1/0$) elicited during the Ultra-Flash profiling task.

Relations between the flavor/taste descriptors ($X = 1/0$ design variables for CATA and Napping, $X$ = intensity/0 for CATA with intensity) and the experimental beers ($Y = 1/0$) were studied by Discriminant-PLSR (D-PLSR; Martens & Martens, 2001) with all variables standardized. Cross validation was performed by splitting the datasets into consecutive segments with eight samples (corresponding to one consumer), thus leaving out one consumer at a time. Martens uncertainty test, a jack-knife based extension of cross validation, was performed to estimate the significance of the model parameters at the optimal number of components, taken at the minimum predicted root mean square error (Martens & Martens, 2000).

At last, data were organized into three separate matrices to allow for comparison of the individual methods sample configurations obtained. First, the frequencies of the CATA descriptors were calculated for each of the eight beers and the counts organized in a table crossing beers and descriptors. For CATA with intensity ratings, a similar table was constructed using averaged attributes ratings. For each Napping sheet, the $X$ and $Y$ coordinates of each sample were determined, using the left bottom corner of the sheet as the origin of the coordinate system. Multiple Factor Analysis (MFA) was used to analyze the data, using the three initial data matrices as individual MFA groups. Multiple Factor Analysis is a multivariate statistical technique which aims at integrating different groups of variables describing the same observations. MFA can be regarded as an enriched PCA and is performed in three steps: first, an initial PCA is computed on each separate group (in our case, CATA, CATA with intensity, and Napping). Secondly, the square root of the first eigenvalue of each individual PCA is extracted and used as scaling factor for each respective data matrix. Finally, the data are re-merged into a global matrix, and a new PCA is performed on this new and "scaled" data. The scores plot represents a consensus map of how the samples were perceived on a global level. Additionally, MFA has the important characteristic that it allows a rapid comparison of the overall configuration with the individual group configurations (i.e. the initial PCAs), both qualitatively, i.e. by visual inspections of the partial points via projection matrix and quantitatively, i.e. by computing regressor vector ($R_V$) coefficients (Robert & Escofier, 1976) to measure configurational congruence. In this study, $R_V$ coefficients were calculated for all possible combination of methodologies for two dimensions and were used to compare sample configurations for each of the three descriptive methods. Mathematically, it can be shown that the $R_V$ coefficient corresponds to the Pearson's correlation coefficient after rearranging the original matrices into vectors (Meyners, 2001). Thus, in the context of the present research high $R_V$ coefficients values would indicate that the methods would yield similar information.

The MFA was carried out using the FactoMineR package (Lê, Josse, & Husson, 2008) in the statistical software R 2.14.1.

## 3. Results

### 3.1. Discriminative ability (A-PLSR)

Data from CATA ($n = 73$), CATA with intensity ($n = 62$) and Napping ($n = 40$) were analyzed by A-PLSR to compare the discriminative ability of the three models. The optimal component number derived from the cross-validation was two for the CATA, three for the CATA with intensity method, and two for the partial Napping. These components accounted respectively for 8%, 11%, and 4% of the total variance in flavors/tastes of the eight beers (Fig. 1). Jack-knife based estimation of the regression coefficients (of the predictor variables, i.e. the samples since they correspond to columns of a design matrix) and visual inspection of the perturbed sub-models revealed that in all three methods all beers were perceived and described differently ($p < 0.05$). In agreement all three methods tended to group beers in two clusters (Fig. 1a–c) containing beers with floral notes (Fynsk Forår, Sea-buckthorn, Thy Pilsner, Pine beer) versus beers with roasted and caramel notes (Stjernebryg, Enebær Stout, Bøgebryg, Valnød Hertug). The configurations differed somewhat between the three methods particularly with regards to the sample correlated with the second PLSR component.

### 3.2. Sensory characterization (D-PLSR)

Data from CATA ($n = 73$), CATA with intensity ($n = 62$) and Napping ($n = 40$) were analyzed by D-PLSR to compare the descriptive ability of the three models in providing a sensory characterization of the samples.

The D-PLSR conducted with the two CATA dataset included 38 flavor descriptors of which respectively 29 and 24 descriptors for CATA and CATA with intensity were found to have a significant effect on the descriptive profile of the eight beers (Table 2). The most frequently used flavor descriptors for CATA and CATA with intensity were hoppy, spicy/herbal, roasted, floral, woody, nutty, berries, caramel and elderflower, words which covered all seven flavor categories. For Napping basic taste descriptors were the most frequently used (e.g. sweet, bitter, sour) followed by the flavor descriptors hoppy, caramel, fruity, citrus and spicy.

The optimal number of component was four for both CATA and CATA with intensity, accounting for 11% and 16% of the validated $Y$ variance respectively. The first PLSR component described the difference in roasted and nutty flavors versus the floral and berry notes whereas the second component explained differences in hoppy versus spicy/herbal and woody flavors as well as floral/berries (only for CATA). The third component explained variation due to woody and spicy versus other characteristics (e.g. hoppy, berries, floral, roasted) for CATA and hoppy versus floral for CATA with intensity (not shown in figure).

For the Napping/UFP method, consumers used a total of approximately 250 words to describe the flavors and tastes of the eight beers. Words mentioned five times or less were kept out of the analysis and synonyms were semantically grouped, following the guidelines given by Perrin et al. (2009). A total of 37 words were retained, of which 26 were found to have significant effect on the description of the eight beers (Table 2). Thirteen Napping descriptors were in common with the CATA/CATA with intensity. Ten of these descriptors were significantly for describing the beers and covered all seven flavor categories whereas the remaining 24 descriptors were unique to the Napping method. The Napping data were optimally described by five PLSR components accounting for 22% of the validated variance (Table 2). The first PLSR component accounted for variation in sweet, liquorice, alcoholic, caramel, full-bodied and strong flavor/tastes versus more sour, fresh, floral, fruity and light notes. In accordance with CATA and CATA with intensity roasted and nutty notes tended to be opposed to floral along the first component of the D-PLSR for the Napping data. The second component described variation due to bitterness versus low bitterness, fruity, floral and sweet characteristics. The third component described variation due to hoppy and malty versus piney flavor notes (figure not shown).

Clear common trends in consumer perception of the beers were observed across all three methods (Fig. 2). For all three descriptive methods the major variance in the beers were caused by differences in roasted, nutty and caramel notes as opposed to floral flavors along the first component. The second component described variation due to hoppy versus spicy/herbal (CATA and CATA with

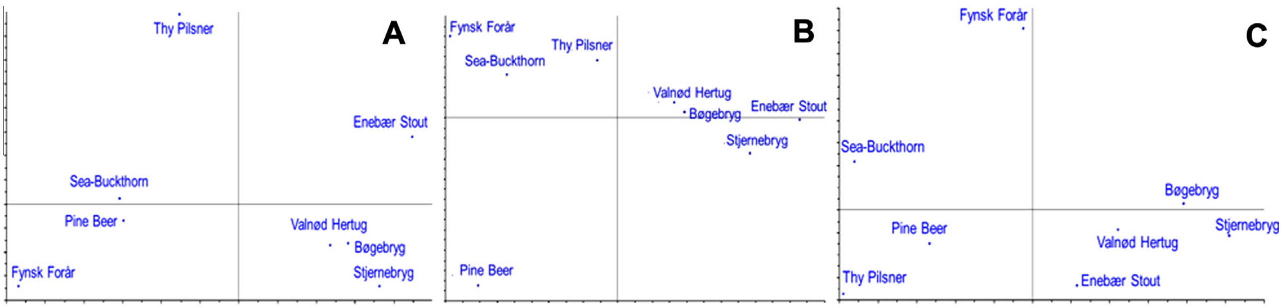


**Fig. 1.** Scores plots from A-PLSR models for CATA (a), CATA with intensity (b), and Napping (c).

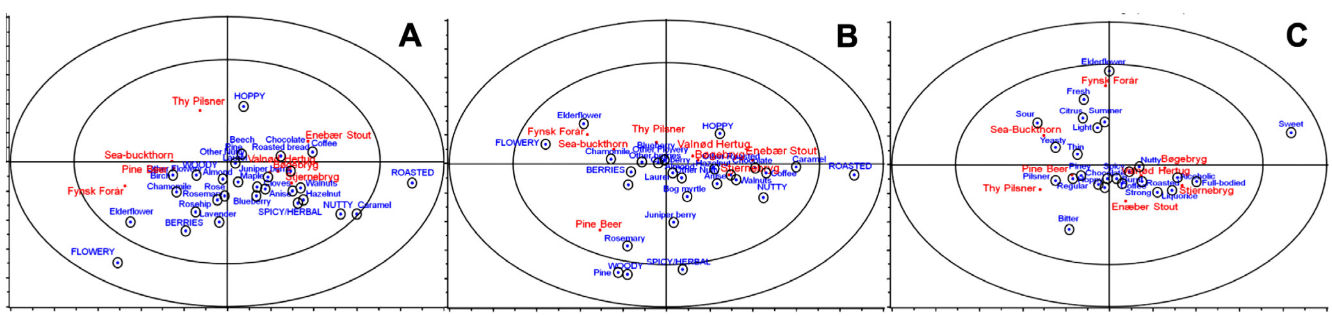


**Fig. 2.** Correlation loadings plot of the D-PLSR models including descriptors for CATA (A), CATA with intensity (B), and Napping (C). Circled descriptors have a significant effect on the sample variation ($p < 0.05$). The inner and outer ellipses represent $R^2 = 50\%$ and $R^2 = 100\%$ respectively.

intensity) or floral and fruity flavors for Napping. Interestingly, the D-PLSR analysis on the Napping dataset showed that the basic tastes, sweet and sour accounted for much of the variation in the first component whereas bitter versus fruity flavors explained variation in the second component. These findings could indicate the additional information could have been gained with CATA/CATA with intensity if basic tastes and the fruity descriptor had been included in the listed descriptors. This also highlights the specific advantage of free descriptors elicitation (such as in Napping/UFP) for uncovering main dimensions relevant for consumers which may be overlooked in methods where subjects are presented with a predefined list of descriptors.

The ability to identify the main sensory characteristics for each of the eight beers (given in Table 1) varied between the three sensory profiling methods. The sea-buckthorn flavored pilsner was described as having sea buckthorn, floral (especially chamomile) and berry flavors using CATA and CATA with intensity whereas Napping revealed that sour was the most descriptive word for the beer (Fig. 2). The pale ale with elderflower (Fynsk Forår) was correctly identified as having a distinct floral and elderflower flavors by all descriptive methods. In contrast, the prominent flavor from the ale with beech twigs (Bøgebryg) and dark ale with walnut (Valnød Hertug) were less clearly identified in their descriptive profile. Bøgebryg and Valnød Hertug tended to be described as having caramel, chocolate, walnut and roasted notes by all methods as well as having nutty and woody notes (Napping, CATA) as well as almond (CATA, CATA intensity) hoppy (CATA intensity), piney (CATA) and to have juniper berry (CATA) and Other berry flavors (CATA intensity). The pine flavored pilsner were perceived to have piney or piney-like flavors (e.g. woody, birch, rosemary) with all three descriptive methods. However, this beer was best differentiated from the other beers with the CATA with intensity method. The anise and liquorice flavors were correctly identified in the strong ale with anise (Stjernebryg) by all descriptive methods and was further characterized as having nutty flavors (all methods) as well as having notes of caramel, cloves, walnut and hazelnut (CATA and CATA with intensity), spicy/herbal flavors (CATA) and as having a sweet taste (Napping). None of the descriptive methods identified the Juniper berries in the Stout with Juniper berries (Enebær Stout); however, this beer was perceived as having distinct coffee and chocolate flavors by all methods. Finally the Thy pilsner was perceived to be hoppy with a hint of floral when using the CATA and CATA intensity method whereas the it was described as a pilsner type beer with bitter and sour tastes when assessed by Napping.

### 3.3. Configurational congruence (MFA)

The last step of our analysis aimed at evaluating the configurational similarity of the sample spaces obtained by the three methods. This was assessed by MFA performed on three cross tabulation matrices containing frequency of concurrencies for each descriptor (CATA), average ratings over samples (CATA with intensity), and the assessors' coordinates (Napping). Fig. 3 below shows the first two dimensions of the consensus MFA sample map (70.5% of the explained variance). The partial configurations obtained by the individual methods are superimposed to the consensus points.

Visual inspection of Fig. 3 shows that all methods provided very similar sample maps. The only partial exception was the Pine beer where the CATA and the CATA with intensity seem to disagree, but only with regards to the variation described by the second MFA component. Accordingly, the three methods correlated similarly with the first MFA component (CATA = 33.7%, CATA with intensity = 33.4%, Napping = 32.9%. Percentages refer to the contribution of individual groups of variables to the MFA component), and differed slightly with regards to the second component (17.5%, 48.9%, and 33.6% respectively).

$R_V$ coefficients were used to compare sample configurations obtained in the three descriptive methods. For all method comparisons the $R_V$ coefficients varied between 0.90 and 0.97, indicating a very high similarity between all three methods. The highest similarity was found between the two CATA methods ($R_V = 0.97$)

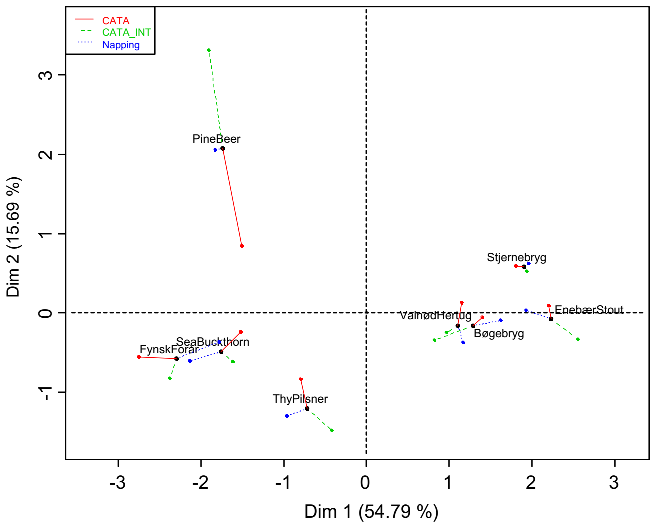


**Fig. 3.** Consensus MFA sample space (first two components) with superimposed partial points from individual methods.

followed by Napping and CATA with intensity ($R_V = 0.93$) whereas the similarity between Napping and the CATA method was slightly lower ($R_V = 0.90$).

## 4. Discussion

In the present study, three consumer-based descriptive methods were compared. Two of the methods were verbal-based – CATA, and a CATA variant with possibility to rate intensity of checked attributes – while one was based on direct sample comparisons (Napping). Despite the different nature, comparative MFA revealed that the overall product map based on data from the three descriptive methods was very similar to the individual beer sample configurations for each method. The high similarity between methods was confirmed by the $R_V$ coefficients ranging from 0.90 to 0.97. These results suggest that the precision and reproducibility of sensory information obtained by consumers with CATA is comparable to that of Napping. This is in concordance with findings by Ares et al. (2010) that projective mapping and CATA questions provide very similar sensory profiles of eight milk desserts, supporting high validity of both sensory methods. The high similarity found between the two CATA methods in the present study indicates that quantitative scaling might not improve the accuracy of the CATA profile, but indicates that a quantitative dimension can be added to CATA without complicating the task for the consumer. The latter point is suggested by the equal time to completion (average of 45 min for both CATA versions), the equal number of checked terms (5.5 on average for the straight CATA and 5.7 for the scaled version), and, indirectly, by the fact that on a separate analysis no significant difference were found on liking ratings between the two CATA groups. A previous comparison between CATA and intensity scaling confirmed a high agreement between the two methods (Parente, Ares, & Manzoni, 2010). Recently Ares, Varela, Rado, and Gimenez (2011a) and Ares, Varela, and Rado (2011b) compared intensity scaling with CATA and projective mapping of orange flavored powdered drinks in one study, and added sorting as a fourth method in a second study. Comparing the data by MFA, the authors also found the methods to provide similar information regarding the sensory profiles, suggesting that stable product configurations can be obtained across different methodologies and different groups of consumers.

The presented work expanded this approach by including A-PLSR and D-PLSR analyses to ascertain additional information. This approach turned out to be very useful in comparing the three descriptive methods. First the discriminative ability of the three methods was analyzed by A-PLSR and comparable overall trends were obtained for the three methods revealing that the eight beers tended to group into clusters of floral beers versus beers with roasted and caramel flavors. Importantly, the three methods were able to discriminate between the beers. CATA with its clearly defined word descriptors might be faster and make it easier for the consumer to discriminate between and describe the beers and therefore improve discriminative ability when compared to Napping. Conversely CATA have previous been criticized for using relatively impoverished dichotomous data, which would supposedly yield a smaller discrimination power. These claims were however not confirmed in the present study.

Next the descriptive ability of the three methods was compared by using D-PLSR, which revealed that the methods tended to vary in both number of words and type of words used.

Comparable overall trends for the three methods was confirmed with D-PLSR, revealing that the eight beers spanned the product space that ranged from beers with roasted, nutty and caramel notes (Stjernebryg, Enebær Stout, Valnød Hertug, Bøgebryg) to beers with floral, berry and piney flavors (Sea buckthorn beer, Fynsk Forår, Thy Pilsner, Pine beer) in one dimension and hoppy (Thy Pilsner) versus spicy/herbal, woody (Pine beer), fruity or floral in the second dimension. The fact that the sweet and sour as well as bitter taste explained much of the variation in the Napping data highlights the importance of the basic tastes in consumer perception of the eight beers and gives additional information that was not captured by the predefined CATA words.

The ability to identify the main sensory characteristics for each of the eight beers was also similar for the three sensory profiling methods. Three ingredients, elderflower in the pale ale (Fynsk Forår), pine in one of the flavoured pilsners, and anise/liquorice in the strong ale (Stjernebryg) were correctly identified by all descriptive methods, as was the Thy pilsner, which was hoppy and described as a pilsner type beer with bitter taste. The dark ale with walnut (Valnød Hertug) tended to be described as having walnut flavors by all methods and was perceived similar to the amber ale with beech twigs (Bøgebryg). The beech flavor in the amber ale was identified with the CATA method whereas the ale only tended to be woody with the Napping method. The sea-buckthorn in the flavored sea buckthorn pilsner was only identified with CATA and CATA with intensity suggesting that novel ingredients might be easier to capture if consumers are influenced by predefined CATA words raising awareness of the novel ingredients. The juniper berries in the Stout (Enebær Stout) were however not identified by any of the descriptive methods despite the help from the CATA words suggesting that some flavors (e.g. novel, unfamiliar, unexpected) might be difficult to capture, and might not be identified unless proper training is conducted.

Even though the ability to identify and describe key sensory attributes in the beers did not vary much between the three methods, small differences were observed. Napping/UFP naturally tended to facilitate the development of a larger and more diverse vocabulary comprised of both sensory attributes and holistic terms whereas CATA is limited to the listed sensory attributes. Conversely, the predefined CATA words aided the description and identification of certain attributes (e.g. woody and sea buckthorn). These findings suggest that the choice of method should be based on the type of descriptive profile required and the resources available. Napping might be beneficial to use when unique, intuitive or creative descriptors are needed (e.g. explorative studies) with a smaller number of consumers, as Napping is slower and more laborious than CATA. CATA might useful in raising awareness of certain attributes and thereby aid the consumer in performing the descriptive profile. CATA is fast and easy and thus suitable for larger groups of consumers. The CATA with intensity variation investigated here did not show significant advantages compared to the "unscaled" version, providing very similar information. However, it should be acknowledged that the samples used in this work were characterized by rather large sensory differences. Scope exists for future research to assess the issue of discriminative power of CATA with less heterogeneous sample sets.

Thus we conclude, in agreement with other authors (Bruzzone, Ares, & Gimenez, 2012), that the number of times consumers checked the presence of an attribute provides already a good estimate of its intensity. Importantly, the CATA with intensity method yielded fewer significant descriptors than the CATA version, possibly due to consumers' inconsistency in the use of scales. This latter conclusion should be considered tentative due to the fact that the sample size in the CATA with intensity condition was slightly lower than in the CATA one. However, a separate analysis treating the ratings in the CATA w/intensity data as normal 1/0 CATA revealed extremely similar results to the scaled version, and returned a significant effect of two variables (lavender and rose) which were found to be non-significant when intensity ratings were used in the analysis (cf. Table 2). This corroborates the idea that potential advantages of scaling might be outweighed by the higher probability of obtaining a noisier dataset.

## 5. Conclusion

In summary, the combination of MFA and PLSR for data analysis of CATA, CATA with intensity and Napping data revealed very good agreement in terms of inter-perceived product differences. The high inter-method reliability shows that consumer data are highly repeatable, and supports the validity of the three methods employed in the study. Additionally, these findings confirm that rapid descriptive methods are suitable to capture differences among products and that they can be a useful tool for capturing and understanding consumer perceptions. The choice of methodology should be based on practical considerations, such as ease of use or whether it is desired to have consumers articulate their own perception of descriptors, or if it sufficient to present them to an existing vocabulary.

## Acknowledgments

Support for this work was provided by the Danish Agency for Science, Technology and Innovation and by the Faculty of Science, University of Copenhagen. The organizing committee of the 5th Eurosense Conference is thanked for the bursary award granted to Davide Giacalone.

## References

Adams, J., Williams, S., Lancaster, B., & Foley, M. (2007). Advantages and uses of check-all-that-apply response compared to traditional scaling of attributes for salty snacks. In *7th Pangborn sensory science symposium*. Minneapolis, MN, USA: Hyatt Regency.

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

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

Ares, G., Varela, P., Rado, G., & Gimenez, A. (2011a). Are consumer profiling techniques equivalent for some product categories? The case of orange-flavored powdered drinks. *International Journal of Food Science and Technology, 46*, 1600–1608.

Bruzzone, F., Ares, G., & Gimenez, A. (2012). Consumers' texture perception of milk desserts. II – Comparison with trained assessors' data. *Journal of Texture Studies, 43*, 214–226.

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., 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.

Dooley, L., Lee, Y., & Meullenet, J. (2010). The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. *Food Quality and Preference, 21*, 394–401.

Ennis, D. M., & Ennis, J. M. (2011). Interpreting applicability scores. *IFP Press, 14*(4), 3–4.

Giacalone, D., Bredie, W. L. P., & Frøst, M. B. (2013). "All-in-one test" (AI1): A rapid and easily applicable approach to consumer product testing. *Food Quality and Preference, 27*, 108–111.

Giacalone, D., Machado Ribeiro, L., & Frøst, M. B. (forthcoming). Consumer-based product profiling: Application of partial Napping for sensory characterization of specialty beers by novices and experts. *Journal of International Food and Agribusiness Marketing* (accepted manuscript).

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213–236.

Lado, J., Vicente, E., Manzoni, A., & Ares, G. (2010). Application of a check-all-that-apply question for the evaluation of strawberry cultivars from a breeding program. *Journal of the Science of Food and Agriculture, 90*, 2268–2275.

Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food: Principles and practices* (2nd ed.). New York, NY: Springer.

Lawless, H. T., Sheng, N., & Knoops, S. S. C. P. (1995). Multidimensional-scaling of sorting data applied to cheese perception. *Food Quality and Preference, 6*, 91–98.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software, 25*, 1–18.

MacFie, H. J., Bratchell, N., Greenhoff, N. K., & Vallis, I. V. (1989). Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests. *Journal of Sensory Studies, 4*, 129–148.

Martens, H., & Martens, M. (2000). Modified jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference, 11*, 5–16.

Martens, H., & Martens, M. (2001). *Multivariate analysis of quality: An introduction* (1st ed.). Chichester, West Sussex, England: John Wiley & Sons Ltd.

Meyners, M. (2001). Permutation tests: Are there differences in product liking? *Food Quality and Preference, 12*, 345–351.

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.

Parente, M. E., Ares, G., & Manzoni, A. V. (2010). Application of two consumer profiling techniques to cosmetic emulsions. *Journal of Sensory Studies, 25*, 685–705.

Perrin, L., & Pages, 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.

Rasinski, K. A., Mingay, D., & Bradburn, N. M. (1994). Do respondents really "mark all that apply" on self-administered questions? *Public Opinion Quarterly, 58*, 400–408.

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

Robert, P., & Escofier, Y. (1976). Unifying tool for linear multivariate statistical-methods- $R_V$-coefficient. *Journal of the Royal Statistical Society Series C –Applied Statistics, 25*, 257–265.

Rosenberg, S., & Kim, M. P. (1975). Method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research, 10*, 489–502.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99–118.

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.