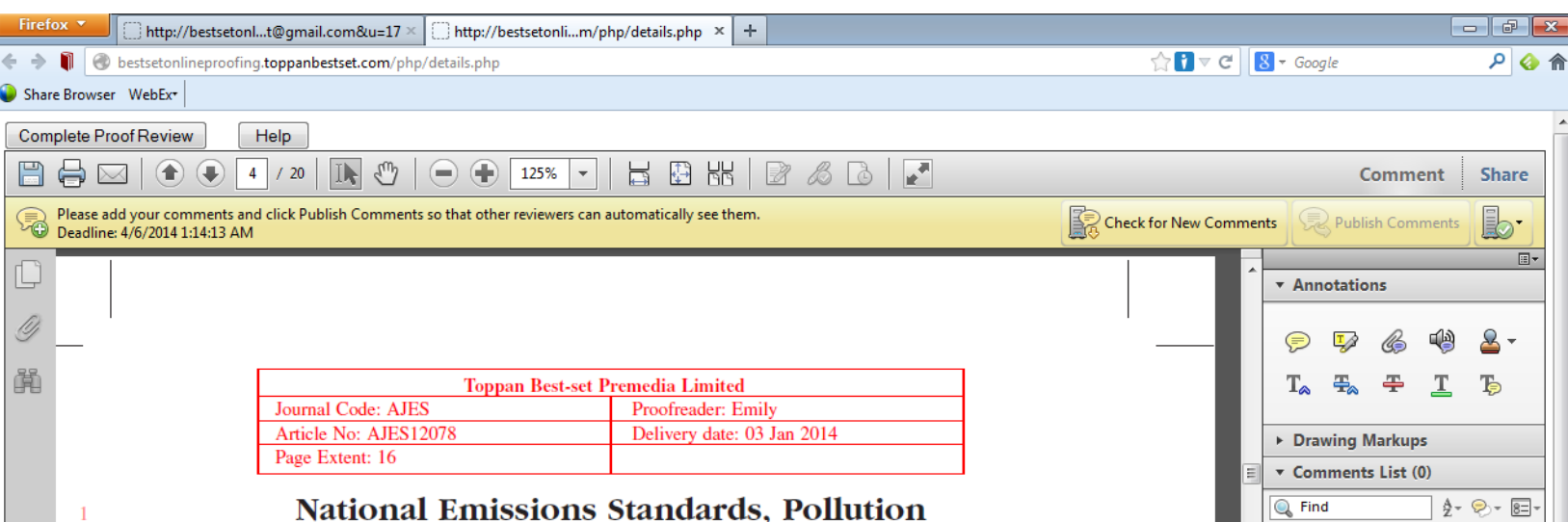


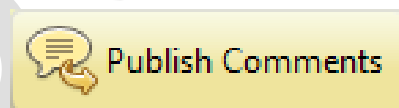
WILEY

Online Proofing System

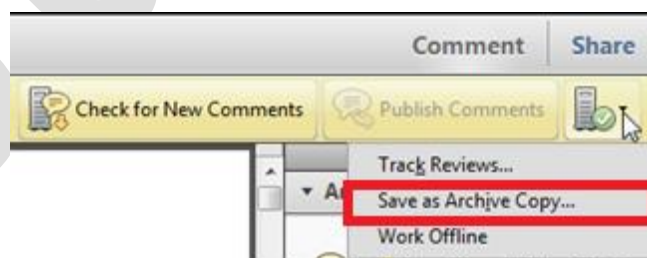
1. Corrections should be marked with the Adobe Annotation & Comment Tools below:



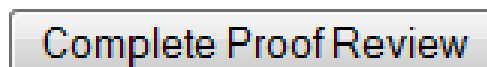
2. To save your proof corrections, click the 'Publish Comments' button. Publishing your comments saves the marked up version of your proof to a centralized location in Wiley's Online Proofing System. Corrections don't have to be marked in one sitting – you can publish corrections and log back in at a later time to add more.



3. When your proof review is complete we recommend you download a copy of your annotated proof for reference in any future correspondence concerning the article before publication. You can do this by clicking on the icon to the right of the 'Publish Comments' button and selecting 'Save as Archive Copy...'.



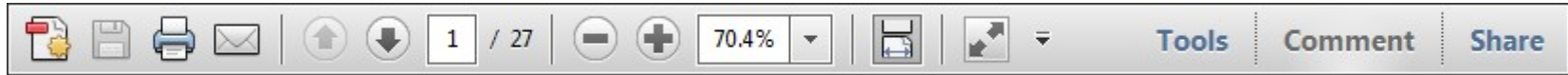
4. When your proof review is complete and you are ready to send corrections to the publisher click the 'Complete Proof Review' button that appears above the proof in your web browser window. **Note: Once you click 'Complete Proof Review' you will not be able to mark any further comments or corrections.**



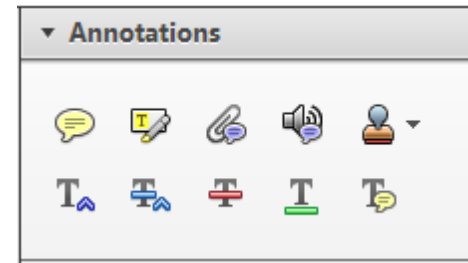
Firefox, Chrome, Safari Users

If your PDF article proof opens in any PDF viewer other than Adobe Reader or Adobe Acrobat, you will not be able to mark corrections and query responses, nor save them. To mark and save corrections, please follow these [instructions to disable the built-in browser PDF viewers in Firefox, Chrome, and Safari](#) so the PDF article proof opens in Adobe within a Firefox or Chrome browser window.

Once you have Acrobat Reader open on your computer, click on the [Comment](#) tab at the right of the toolbar:



This will open up a panel down the right side of the document. The majority of tools you will use for annotating your proof will be in the [Annotations](#) section, pictured opposite. We've picked out some of these tools below:



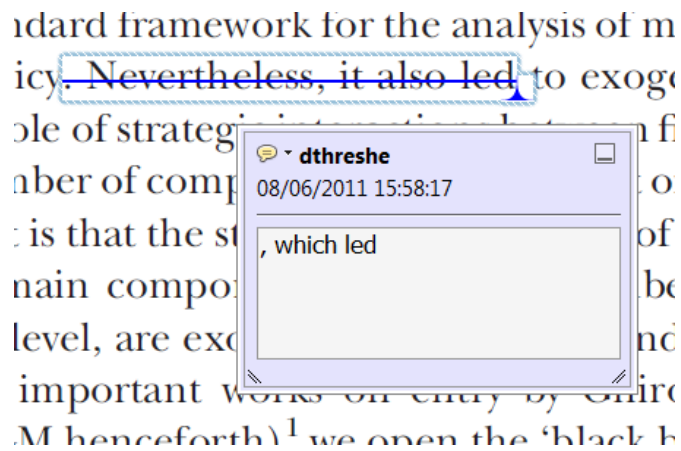
1. Replace (Ins) Tool – for replacing text.



Strikes a line through text and opens up a text box where replacement text can be entered.

How to use it

- Highlight a word or sentence.
- Click on the [Replace \(Ins\)](#) icon in the Annotations section.
- Type the replacement text into the blue box that appears.



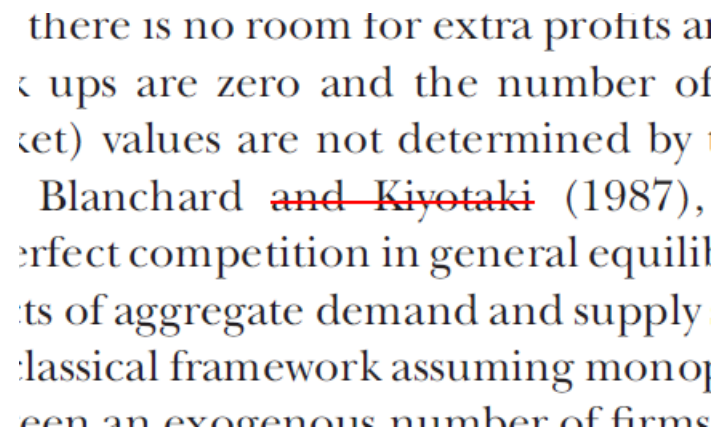
2. Strikethrough (Del) Tool – for deleting text.



Strikes a red line through text that is to be deleted.

How to use it

- Highlight a word or sentence.
- Click on the [Strikethrough \(Del\)](#) icon in the Annotations section.



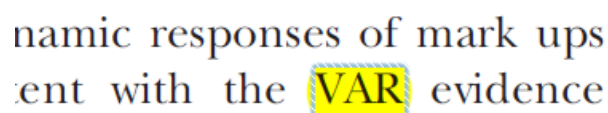
3. Add note to text Tool – for highlighting a section to be changed to bold or italic.



Highlights text in yellow and opens up a text box where comments can be entered.

How to use it

- Highlight the relevant section of text.
- Click on the [Add note to text](#) icon in the Annotations section.
- Type instruction on what should be changed regarding the text into the yellow box that appears.



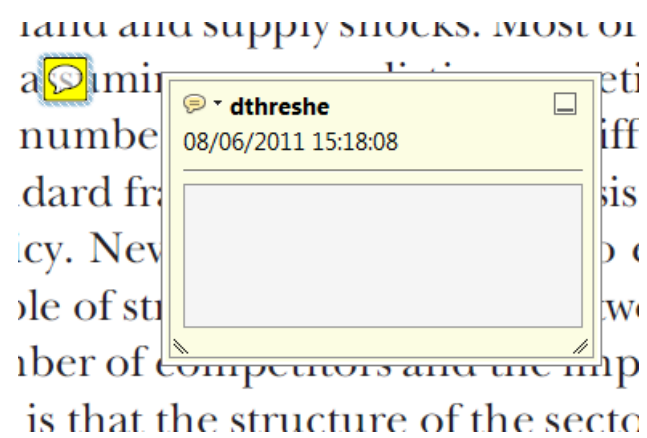
4. Add sticky note Tool – for making notes at specific points in the text.



Marks a point in the proof where a comment needs to be highlighted.

How to use it

- Click on the [Add sticky note](#) icon in the Annotations section.
- Click at the point in the proof where the comment should be inserted.
- Type the comment into the yellow box that appears.



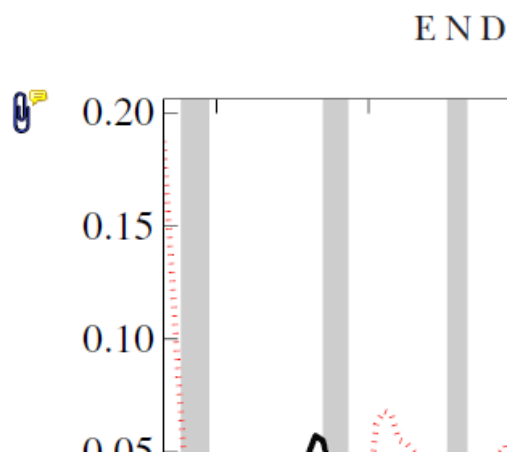
5. Attach File Tool – for inserting large amounts of text or replacement figures.



Inserts an icon linking to the attached file in the appropriate place in the text.

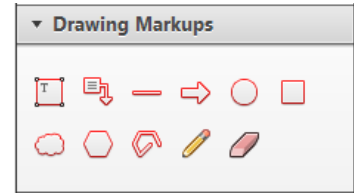
How to use it

- Click on the [Attach File](#) icon in the Annotations section.
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.



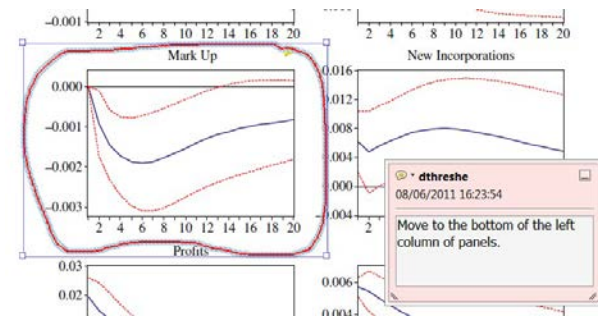
6. Drawing Markups Tools – for drawing shapes, lines and freeform annotations on proofs and commenting on these marks.

Allows shapes, lines and freeform annotations to be drawn on proofs and for comment to be made on these marks.



How to use it

- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, move the cursor over the shape until an arrowhead appears.
- Double click on the shape and type any text in the red box that appears.



| | |
|-----------------------|----------------------------|
| Journal Code: JOSS | Proofreader: Mony |
| Article No: JOSS12135 | Delivery date: 14 Jan 2015 |
| Page Extent: 10 | |

EVALUATION OF DATA AGGREGATION IN POLARIZED SENSORY POSITIONING

LUCÍA ANTÚNEZ^{1,4}, ANA SALVADOR², LUIS de SALDAMANDO¹, PAULA VARELA³, ANA GIMÉNEZ¹ and GASTÓN ARES¹

¹Departamento de Ciencia y Tecnología de Alimentos, Facultad de Química, Universidad de la República, Montevideo, Uruguay
²Instituto de Agroquímica y Tecnología de Alimentos (CSIC), Paterna, Valencia, Spain
³Nofima AS, Ås, Norway

⁴Corresponding author.
TEL: +59829248003;
FAX: +59829241906;
EMAIL: lantunez@fq.edu.uy

Accepted for Publication December 18, 2014

doi:10.1111/joss.12135

ABSTRACT

The aim of the present work was to evaluate data aggregation when using two polarized sensory positioning (PSP) approaches for sensory characterization with consumers. Two consumer studies with different product categories (orange-flavored powdered drinks and chocolate milk beverages) were carried out. In each study two PSP approaches were considered: PSP with scales and triadic PSP (t-PSP). For each approach, one-third of the consumers evaluated the whole sample set, whereas the other two-thirds evaluated the sample set split in two subsets. Results showed that sample configurations for the evaluation of the whole and the split set by different consumer groups were relatively well correlated (RV coefficients higher than 0.79). However, agreement between the configurations differed between the studies, which can be explained by the degree of difference among samples. Besides, differences in consumers' dissimilarity scores and conclusions regarding similarities and differences among samples were identified when comparing both data sets (with and without data aggregation). Regarding the comparison of the two PSP approaches, in the two studies better agreement between sample configurations was obtained for t-PSP. However, in one of the studies PSP with scales provided better results for the evaluation of a repeated sample by different consumer groups.

PRACTICAL APPLICATIONS

Polarized sensory positioning has been gaining popularity in the last years. The main advantage of this methodology over other holistic methodologies is that it allows aggregating data from different studies, which is particularly interesting when working with consumer-based sensory characterization. Results from the present work showed that aggregation of data from the evaluation of split sample sets by different consumer groups provided similar results than the evaluation of the whole sample set. However, conclusions regarding similarities and differences among samples differed in one of the studies, which suggests that care must be taken when aggregating data from the evaluation of similar samples by different consumer groups.

INTRODUCTION

Descriptive analysis with trained assessors is one of the most extensively used methodologies for sensory product characterization (Stone *et al.* 1974; Meilgaard *et al.* 1999; Murray *et al.* 2001). In this methodology, assessors are extensively trained in attribute recognition and scaling

using clearly defined references (Lawless and Heymann 2010). For this reason, descriptive analysis provides detailed, accurate, reproducible and stable time results. However, training and maintaining a sensory panel can be time consuming and expensive, which makes descriptive analysis difficult to apply in many situations (Murray *et al.* 2001; Varela and Ares 2012). Therefore, interest in the development of

1 less sophisticated and faster methodologies has markedly
2 grown in the last decade (Valentin *et al.* 2012; Varela and
3 Ares 2012).

4 Holistic methodologies, such as sorting and projective
5 mapping, are one of the novel approaches for sensory char-
6 acterization (Valentin *et al.* 2012). They are based on the
7 evaluation of global differences among samples, enabling
8 identification of the main sensory characteristics respon-
9 sible for perceived similarity among samples (Ares and
10 Varela 2014). Despite the fact that these methodologies have
11 been reported to provide valid and reliable information
12 [3] (Risvik *et al.* 1997; Chollet and Valentin 2001; Faye *et al.*
13 2004; Chollet *et al.* 2011; Dehlholm *et al.* 2012; Hopfer and
14 [4] Heymann 2013), one of their main disadvantages is that the
15 entire set of products must be simultaneously evaluated in
16 the same session (Teillet *et al.* 2010; Varela and Ares 2012;
17 Ares *et al.* 2013). This restricts the number of samples that
18 can be evaluated as well as the possibility of comparing
19 samples evaluated in different moments in time.

20 In order to overcome this limitation of holistic method-
21 ologies, Teillet *et al.* (2010) have developed polarized
22 sensory positioning (PSP). This methodology is based on
23 the comparison of samples with a fixed set of reference
24 products called “poles” (Teillet 2014). Despite the fact that
25 this methodology was originally developed to explore the
26 sensory characteristic of water with trained assessors (Teillet
27 *et al.* 2010), it has been successfully used with naïve con-
28 sumers (Ares *et al.* 2013; de Saldamando *et al.* 2013; Teillet
29 2014).

30 Two main PSP approaches have been reported: PSP with
31 scales and triadic PSP (t-PSP). In PSP with scales assessors
32 have to use unstructured scales to rate the overall similarity
33 (or dissimilarity) between each sample and each one of the
34 poles. Teillet *et al.* (2010) used unstructured scales ranging
35 from “exactly the same taste” to “totally different taste”
36 when evaluating the taste of mineral waters. t-PSP can be
37 regarded as similar to a “polarized triad” test (MacRae *et al.*
38 1990) in which similarities and dissimilarities to poles are
39 estimated from co-occurrences. Assessors are asked to indi-
40 cate to which of the poles a sample is more similar and to
41 [5] which it resembles the least (Teillet *et al.* 2014).

42 Regardless of the specific approach considered for sample
43 evaluation, PSP approaches are based on the comparison of
44 samples with a set of poles, which are kept constant across
45 different sessions. Therefore, PSP makes it possible to aggre-
46 gate data from different sessions and to accumulate data
47 from different studies. Considering the increasing interest in
48 consumer-based sensory characterizations and the difficul-
49 ties usually encountered for recruiting consumers for repli-
50 cated evaluations, the possibility of aggregating data from
51 sensory characterization studies performed with different
52 consumers is particularly interesting in both industrial and
53 academic applications. However, to the authors’ knowledge

54 no study evaluating data aggregation from sensory charac-
55 terization studies with consumers using PSP has been pub-
56 lished in refereed journals.

57 In this context, the aim of the present work was to evalu-
58 ate data aggregation when using two PSP approaches (PSP
59 with scales and t-PSP) for sensory characterization with
60 consumers. Sample configurations obtained when different
61 consumer groups evaluated the whole and split sample sets
62 using PSP with scales and t-PSP were compared in two
63 studies with two different product categories.

64 MATERIALS AND METHODS 65

66 Two studies with two different product categories (orange-
67 flavored powdered drinks and chocolate milk beverages)
68 with 240 consumers were conducted to evaluate data aggre-
69 gation in PSP. In each study two PSP approaches were con-
70 sidered: PSP with scales and t-PSP. For each PSP approach a
71 between-subject design was used to compare sample con-
72 figurations for the evaluation of the whole and the split
73 sample sets.

74 Samples 75

76 A total of seven samples of commercial orange-flavored
77 powdered drinks were used in study 1. All samples were
78 available in the Uruguayan market and were purchased
79 from local supermarkets in Montevideo (Uruguay). The set
80 involved six samples (A–F) and a set of three poles (PA, PB
81 and PC). Two of the poles were identical to the samples in
82 order to evaluate the validity of the methodology. Poles
83 were selected based on results from a previous study that
84 used a projective mapping methodology to identify the
85 sensory characteristics responsible for the main differences
86 among commercial samples of orange-flavored powdered
87 drinks (Ares *et al.* 2013). Three main groups of samples
88 were identified in that study: one is characterized by its low
89 total flavor intensity, a second one by its sourness and a
90 third group was described as sweet and with intense orange
91 flavor. Considering these results one pole from each group
92 was selected: low total flavor intensity (PB), one sweet drink
93 with intense orange flavor, formulated with sugar (PC) and
94 the other characterized by its sourness formulated with low-
95 calorie sweeteners (PA) (Ares *et al.* 2013). Table 1 provides a
96 description of the samples in terms of their main charac-
97 teristics and market positioning. Samples were prepared fol-
98 lowing the recommendations provided by the manufacturer
99 on the package. The powders were diluted in tap water and
100 stored in a fridge at 10C until they were served to consum-
101 ers within 4 h.

102 Study 2 was carried out with samples of commercial
103 chocolate milk beverages available in the Spanish market.
104 The set involved a total of seven samples (G–M). Three
105

TABLE 1. DESCRIPTION OF THE SEVEN ORANGE-FLAVORED POWDERED DRINKS EVALUATED IN STUDY 1 IN TERMS OF MARKET SEGMENT AND MAIN CHARACTERISTICS

| Sample | Market segment | Main characteristics |
|--------|----------------|--|
| A | Premium | Contains sugar and vitamins A, C, B2, B3, B6, folic acid |
| B | Premium | Without sugar |
| C | Economy | Contains sugar and sweeteners |
| D | Economy | Contains sugar and sweeteners |
| E, PB | Economy | Contains sugar and sweeteners |
| F, PC | Medium | Contains sugar and vitamins A, C, B2, B3, B6 and B9 |
| PA | Premium | Without sugar |

PA, PB and PC refer to the poles used in the evaluation.

poles (A, B and C) were selected based on results from a preliminary projective mapping study in which 20 consumers evaluated eight samples of commercial chocolate milk beverages. Poles represented the main sensory characteristics responsible for the similarities and differences among samples. As in study 1, two of the poles were identical to the samples. A description of the main characteristics of the samples is provided in Table 2. Chocolate milk beverages were stored in a fridge at 10C until they were served to consumers.

Participants

Study 1 was carried out with 240 consumers (ages ranging from 18 to 57 years old, 68% female and 32% men). All of them were recruited from the consumer database of the Food Science and Technology Department of Universidad de la República (Uruguay) based on their availability and interest to participate.

A total of 240 consumers participated in study 2 (age ranging from 18 to 69 years old, 60% female and 40% men). Consumers were recruited from the university campus

TABLE 2. DESCRIPTION OF THE CHOCOLATE MILK BEVERAGES INCLUDED IN STUDY 2

| Sample | Main characteristics |
|--------|--|
| G | Contains skimmed milk; cocoa (1.4%); milk powder; vitamins E, A and D; vanilla aroma |
| H | Contains milk, whey, cocoa (1.2%) |
| I | Contains skimmed milk, cocoa (1%), without lactose |
| J, PA | Contains skimmed milk, cocoa (1.2%) |
| K | Contains skimmed milk, dietary fiber, cocoa (1.5%), sweetener, vitamins A and D |
| L | Contains soybeans, cocoa (1%), vitamins B2, B12 and D2 |
| M, PB | Contains milk, cocoa (0.9%) |
| PC | Contains skimmed milk, whey, cocoa (1.2%) |

PA, PB and PC refer to the poles used in the evaluation.

(Universidad Politécnica, Valencia, España) and from Instituto de Agroquímica y Tecnología de Alimentos (Valencia, Spain) based on their availability and interest to participate in the study.

Data Collection

In each study consumers were randomly divided into two groups of 120, each of which ~~was subdivided into three groups. Each group~~ performed a different task. Group 1 evaluated the samples using PSP with scales, whereas group 2 evaluated samples using t-PSP. Besides, each consumer group was subdivided into three groups of 40, each of which evaluated a different sample set. Subgroup A evaluated the whole sample set, whereas subgroups B and C evaluated a split set. Split sets consisted of splitting the samples in two sets (set A and set B). In study 1, each subgroup evaluated three samples, whereas in study 2 the split sets were composed of four samples with one repeated sample (sample G). A summary of the studies is shown in Table 3.

The procedure for data collection in study 1 and 2 was the same. Consumers received 60 mL of each of the three poles and approximately 30 mL of the different samples, which were served in plastic glasses coded with three-digit random numbers. The order in which participants received samples differed among participants, following a design balanced for order and carry-over effects (William's Latin square). Assessors were told that they had to complete the study according to their own criteria taking into account that there were no right or wrong answers. Mineral water was available for rinsing between samples. Testing took place in a sensory laboratory in standard sensory booths designed in accordance with ISO 8589 (ISO 2007) under artificial daylight and temperature control (22C).

PSP with Scales. In the tasks involving PSP with scales, assessors were asked to rate the overall difference between each sample and each one of the poles using an unstructured scale anchored from "exactly the same" to "totally different."

TABLE 3. DESCRIPTION OF TASKS PERFORMED BY EACH CONSUMER SUBGROUP (N = 40) IN STUDY 1 AND STUDY 2 FOR EVALUATING WHOLE AND SPLIT SETS USING POLARIZED SENSORY POSITIONING (PSP) WITH SCALES AND TRIADIC POLARIZED SENSORY POSITIONING

| Group | Subgroup | Methodology | Sample set | Number of samples |
|-------|----------|-----------------|-------------|-------------------|
| 1 | A | PSP with scales | Whole set | 6 |
| | B | | Split set A | 3 |
| | C | | Split set B | 3 |
| 2 | A | Triadic PSP | Whole set | 7 |
| | B | | Split set A | 4 |
| | C | | Split set B | 4 |

1 **t-PSP.** In t-PSP tasks, consumers were asked to indicate to
2 which pole each of the samples resembled the most and to
3 which pole it resembled the least.

4 **Data Analysis**

5 **PSP with Scales.** Data from PSP with scales were consid-
6 ered as sensory descriptors and consequently analyzed using
7 principal component analysis (PCA) (Teillet 2014). For each
8 sample, the average score was calculated and a matrix con-
9 taining samples in rows and poles in columns was con-
10 structed. PCA was applied on the correlation matrix of
11 average scores. When samples were evaluated by different
12 groups of consumers, data were analyzed by binding the
13 matrices obtained for each consumer group.

14 In study 2, one of the repeated samples, selected at
15 random, was considered as supplementary individual in the
16 analysis (sample G*).

17 Analysis of variance (ANOVA) was carried out to identify
18 significant differences in the difference scores between
19 samples and each of the poles between the evaluations with
20 and without data aggregation. Type of evaluation, sample
21 and their interaction were considered as fixed sources of
22 variance. A significance level of 5% was considered. When
23 the effects were significant, honestly significant differences
24 were calculated using Tukey's test.

25 **t-PSP.** Data from t-PSP were analyzed considering the pole
26 to which the sample resemble the most (named A+, B+ or
27 C+) and the pole to which the sample resemble the least
28 (named A-, B- and C-) as qualitative variables. A fre-
29 quency table containing the number of times a sample was
30 regarded as most similar and most different to each of the
31 poles was constructed and analyzed by means of correspon-
32 dence analysis (CA) (Teillet *et al.* 2014). When samples were
33 evaluated by different groups of consumers, data were ana-
34 lyzed by binding the frequency tables obtained for each
35 consumer group.

36 In study 2, one of the repeated samples, selected at
37 random, was considered as supplementary individual in the
38 analysis (sample G*).

39 **Comparison of Sample Configurations.** The RV coef-
40 ficient (Robert and Escoufier 1976) was used to evaluate the
41 agreement between the first two dimensions of sample con-
42 figurations obtained from the evaluation of the whole and
43 the split sample sets, as well as the similarity between
44 sample configurations obtained using t-PSP and PSP with
45 scales. The significance of the RV coefficient was tested
46 using a permutation test (Josse *et al.* 2008).

47 All statistical analyses were performed with R language
48 (R Core Team 2013) using the package FactoMineR (Lê
49 *et al.* 2008).

54 **RESULTS**

55 **Study 1**

56 As shown in Figs. 1 and 2, regardless of the PSP approach
57 used for evaluating samples, the percentage of variance
58 explained by the first and second dimensions of the
59 PCA/CA did not largely differ between the evaluation of the
60 whole set and data aggregation from the evaluation of
61 the split set by different consumer groups.

62 For both methodologies, sample configurations obtained
63 through data aggregation from the evaluation of the split set
64 by different consumer groups was similar to those obtained
65 from the evaluation of the whole set. The RV coefficients
66 between sample configurations were significant and higher
67 than 0.90 (Table 4). Besides, for both PSP with scales and
68 t-PSP the position of the samples with respect to the poles
69 was similar when the configuration was based on data from
70 the evaluation of the whole set or based on data aggre-
71 gation from the evaluation of the split set (Figs. 1 and 2,
72 respectively).

73 The validity of the methodology was evaluated consider-
74 ing the evaluation of two blind samples (E and F) identical
75 to two of the poles (PB and PC, respectively). As shown in
76 Fig. 1, in PSP with scales samples E and F were located
77 opposite to the direction of increasing difference with poles
78 PB and PC, respectively, regardless of the type of evaluation
79 (whole set or data aggregation from the evaluation of the
80 split set by different consumer groups). Meanwhile, in t-PSP
81 samples E and F were located close to the columns that
82 represent similarity to poles B and C (PB+ and PC+),
83 respectively, in both evaluations (Fig. 2).

84 Despite the high similarity between sample configura-
85 tions, some differences in conclusions regarding similarities
86 and differences among samples were identified. For PSP
87 with scales, the relative position of sample D in the sensory
88 space markedly differed between sample configurations
89 obtained through the evaluation of the whole set and the
90 data aggregation from the evaluation of the split set by dif-
91 ferent consumer groups. When the whole set was evaluated,
92 sample D was located in a distinct position (Fig. 1A), while
93 when sample configurations obtained by aggregating data
94 from the evaluation of the split set was considered, sample
95 D was regarded as similar to samples E and C (Fig. 1B).

96 ANOVA was used to assess if dissimilarity scores obtained
97 by the evaluation of the whole sample set and data aggre-
98 gation from the evaluation of the split set by different con-
99 sumer groups significantly differed. As shown in Table 5,
100 difference scores between samples and poles A and C were
101 significantly affected by the type of evaluation (whole
102 sample set or data aggregation from the evaluation of the
103 split set). On average, difference scores were higher when
104 consumers evaluated the whole sample set than when
105

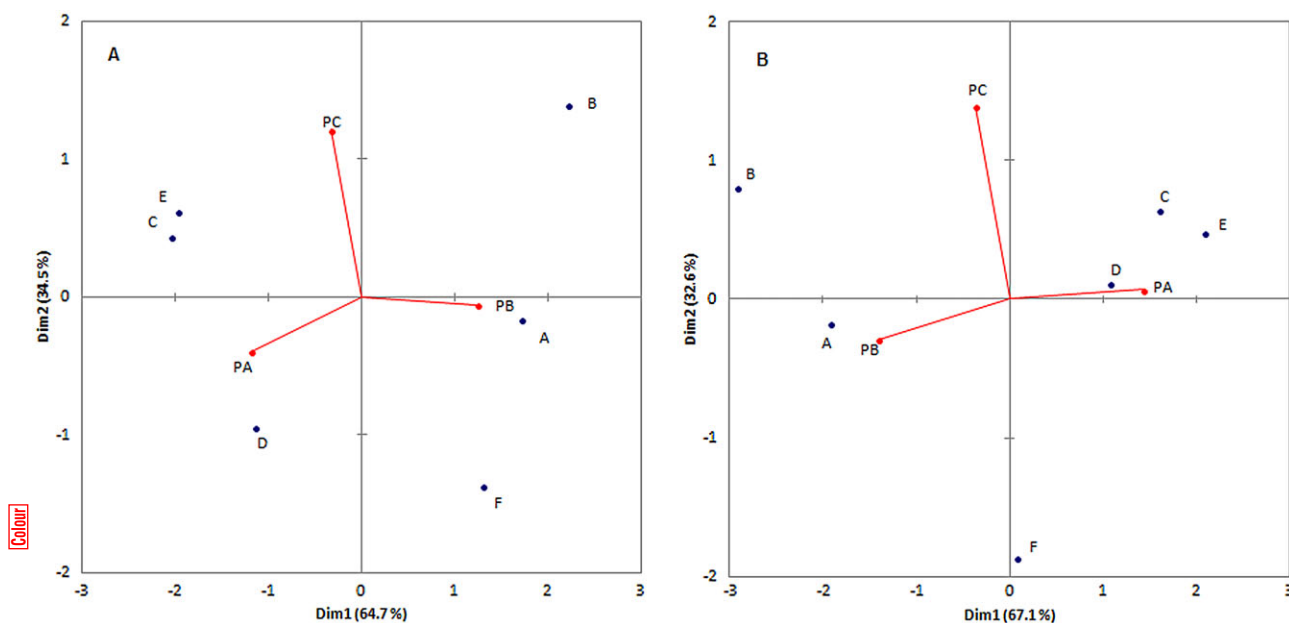


FIG. 1. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF PRINCIPAL COMPONENT ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING POLARIZED SENSORY POSITIONING WITH SCALES FOR THE EVALUATION OF ORANGE-FLAVORED POWDERED DRINKS
 Sample E was identical to pole B (PB) and sample F identical to pole C (PC).

different groups evaluated the split set. Type of evaluation did not significantly affect difference scores between samples and pole B. Despite the fact that type of evaluation significantly affected difference scores, the interaction

between sample and type of evaluation was not significant for the evaluation of poles B and C (Table 5). However, the interaction was significant for the evaluation of pole A. This suggests that the type of evaluation significantly affected

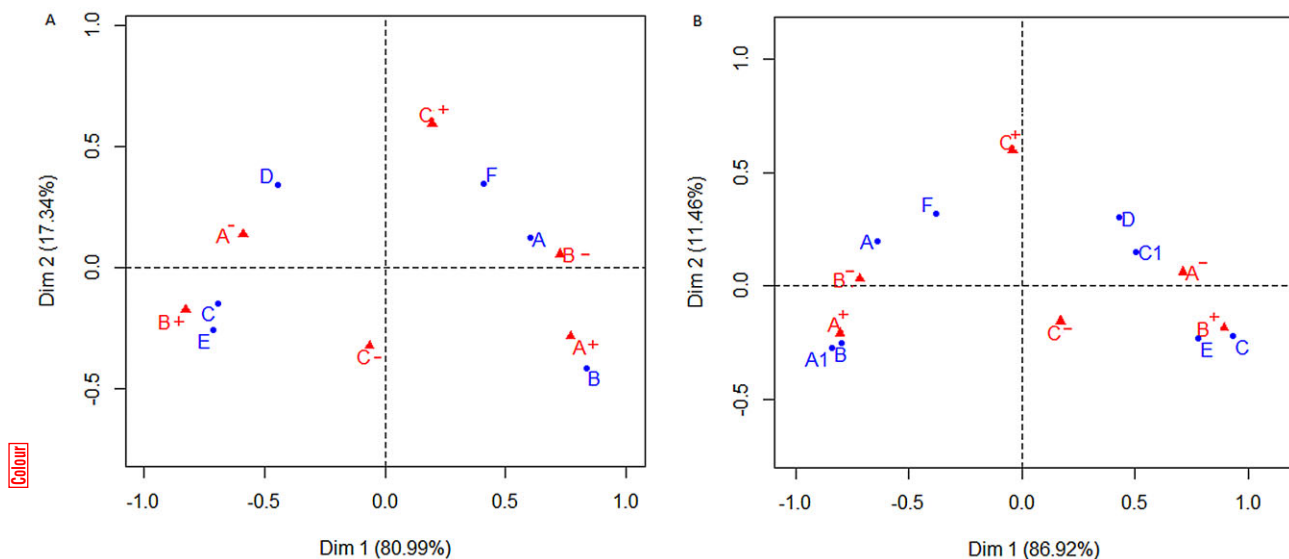


FIG. 2. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF CORRESPONDENCE ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING TRIADIC POLARIZED SENSORY POSITIONING FOR THE EVALUATION OF ORANGE-FLAVORED POWDERED DRINKS
 Sample E was identical to pole B (PB) and sample F identical to pole C (PC).

| Study | Comparison | PSP approach | RV coefficient | P value |
|---------|--|------------------------------|----------------|---------|
| Study 1 | Whole set versus data aggregation from the split set | PSP with scales | 0.92 | 0.006 |
| | | t-PSP | 0.98 | 0.014 |
| Study 2 | Whole set versus data aggregation from the split set | t-PSP versus PSP with scales | 0.96 | 0.010 |
| | | PSP with scales | 0.79 | 0.017 |
| | | t-PSP | 0.91 | 0.012 |
| | Whole set | t-PSP versus PSP with scales | 0.44 | 0.125 |

The RV coefficient between sample configurations obtained from the evaluation of the whole sample set using PSP and t-PSP is also included.

TABLE 4. RV COEFFICIENT BETWEEN THE FIRST TWO DIMENSIONS OF SAMPLE CONFIGURATIONS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET AND DATA AGGREGATION FROM THE EVALUATION OF THE SPLIT SET BY DIFFERENT CONSUMER GROUPS USING POLARIZED SENSORY POSITIONING (PSP) AND TRIADIC PSP (T-PSP) IN STUDY 1 AND STUDY 2

how consumers evaluated difference between samples and one of the poles (pole A), which could affect sample configurations. Despite this significant effect on the evaluation of the degree of difference between samples and pole A, sample configurations obtained by the evaluation of the whole set and data aggregation of the evaluation of the split set did not largely differ (Fig. 1).

Regarding t-PSP, the main difference between sample configurations was related to the relative positioning of sample C. When the whole set was evaluated, this sample was perceived as similar to sample E (Fig. 2A), whereas when sample configuration obtained by aggregating data from the evaluation of the split set with different consumer groups was taken into account, this sample was located closer to sample D than to sample E (Fig. 2B).

Sample configurations were not largely affected by the PSP approach used by consumers to evaluate samples. The RV coefficient between sample configurations from PSP with scales and t-PSP tasks was 0.96 (Table 4).

Study 2

The percentage of variance/inertia explained by the first two dimensions of the PCA/CA did not largely differ between the evaluation of the whole set and data aggregation from the evaluation of the split set by different consumer groups for both PSP with scales (Fig. 3) and t-PSP (Fig. 4), respectively.

As shown in Table 4, when t-PSP was considered the RV coefficient between sample configurations obtained by the evaluation of the whole set and aggregated data from the evaluation of the split set by different consumer groups was higher than 0.90, indicating good agreement. However, when PSP with scales was used the RV coefficient between sample configurations was significant but markedly lower (Table 4).

When PSP with scales was used several differences in conclusions regarding similarities and differences between the evaluation of the whole set and aggregated data from the evaluation of the split set were observed. As shown in Fig. 3, the position of samples K and G markedly differed between sample configurations. When the whole set was evaluated, samples K and G were located close to sample M and far from sample J (Fig. 3A), while they were located in the opposite relative position when the split set was evaluated by two consumer groups (Fig. 3B). The projection of sample G* (considered as supplementary individual on the sensory space) was located close to sample G.

Furthermore, conclusions regarding the degree of similarity between samples and poles A and B were similar regardless of the type of evaluation (Fig. 3). However, the evaluation of the degree of similarity between samples and pole C changed. As shown in Fig. 2, dissimilarity scores with respect to pole C were highly correlated to dissimilarity scores with respect to pole B when consumers evaluated the whole sample set, whereas they were correlated to

| Study | Source of variation | Pole | | |
|---------|---------------------------|------------------------|------------------------|------------------------|
| | | A | B | C |
| Study 1 | Sample | 64.39 ($P < 0.0001$) | 68.8 ($P < 0.0001$) | 11.05 ($P < 0.0001$) |
| | Type of evaluation | 4.47 ($P = 0.0350$) | 0.42 ($P = 0.5155$) | 5.8 ($P = 0.0165$) |
| | Sample*Type of evaluation | 3.53 ($P = 0.0038$) | 0.83 ($P = 0.531$) | 0.8 ($P = 0.5484$) |
| Study 2 | Sample | 12.96 ($P < 0.0001$) | 41.07 ($P < 0.0001$) | 8.89 ($P < 0.0001$) |
| | Type of evaluation | 1.26 ($P = 0.2627$) | 0.22 ($P = 0.6396$) | 0.34 ($P = 0.5594$) |
| | Sample*Type of evaluation | 1.45 ($P = 0.1925$) | 1.00 ($P = 0.4267$) | 1.59 ($P = 0.1470$) |

TABLE 5. F AND P VALUES (BETWEEN BRACKETS) FROM THE ANALYSIS OF VARIANCE PERFORMED ON DISSIMILARITY SCORES BETWEEN SAMPLE AND THE POLES (A, B AND C) OBTAINED FROM THE EVALUATION OF THE WHOLE AND SPLIT SAMPLE SETS USING POLARIZED SENSORY POSITIONING IN STUDY 1 AND STUDY 2

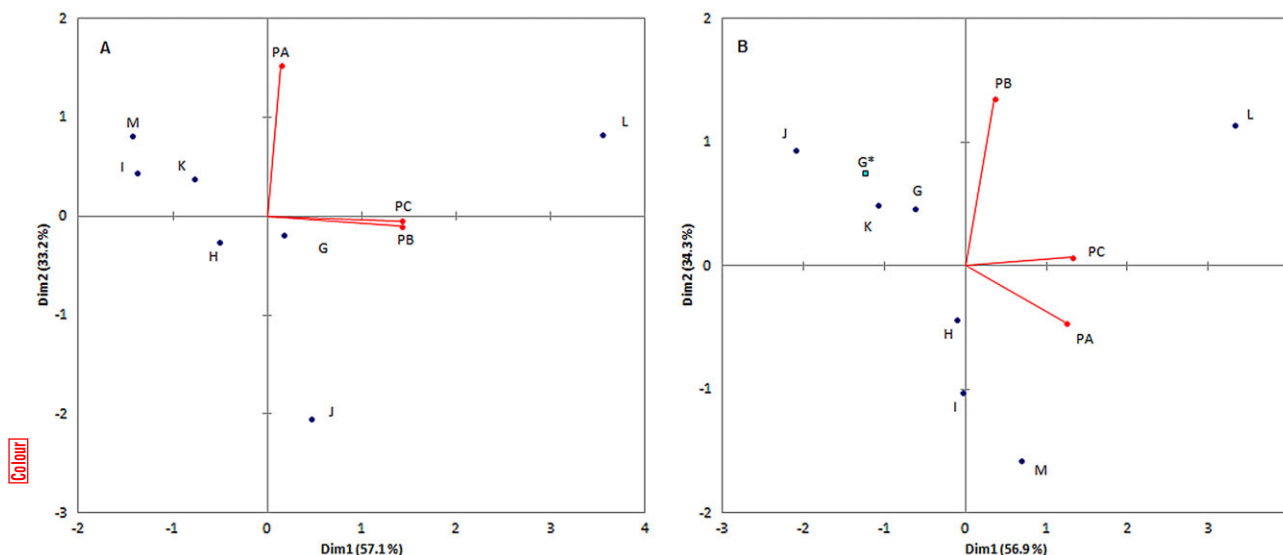


FIG. 3. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF PRINCIPAL COMPONENT ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING POLARIZED SENSORY POSITIONING WITH SCALES FOR THE EVALUATION OF CHOCOLATE MILK BEVERAGES. Sample J was identical to pole A (PA) and sample M identical to pole B (PB).

dissimilarity scores with respect to sample A when aggregated data from the evaluation of the split set by different consumer groups were considered.

As shown in Table 5, average overall difference scores between samples and poles were not significantly affected by

the type of evaluation (whole versus split set) or the interaction between samples and type of evaluation.

When t-PSP was considered, differences between sample configurations from the evaluation of the whole set and the aggregated data from the different consumer groups were

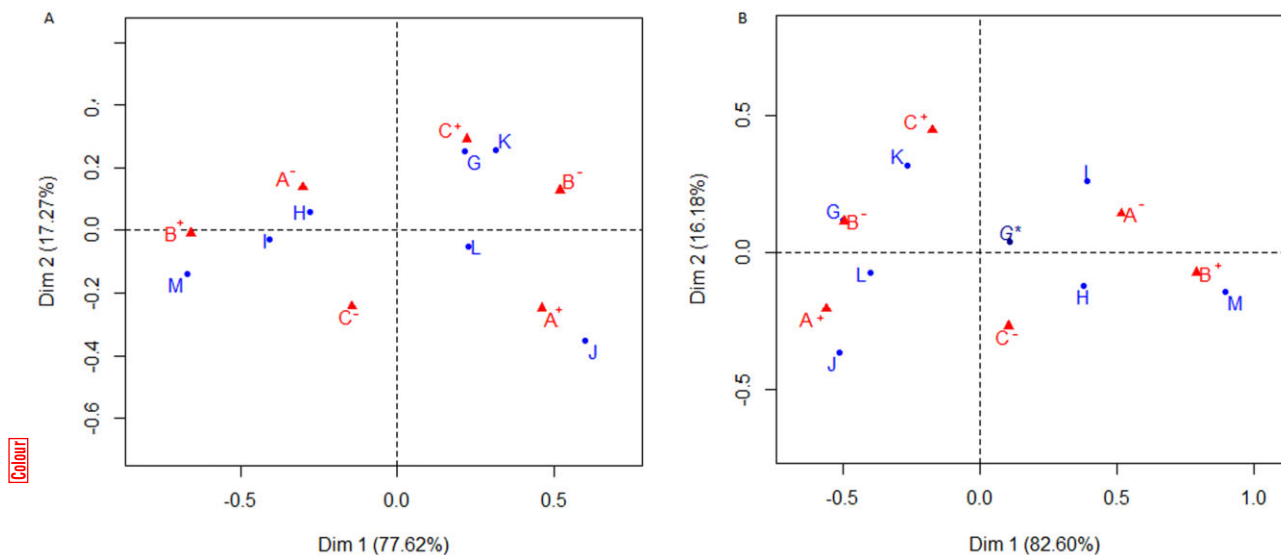


FIG. 4. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF CORRESPONDENCE ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING TRIADIC POLARIZED SENSORY POSITIONING FOR THE EVALUATION OF CHOCOLATE MILK BEVERAGES. Sample J was identical to pole A (PA) and sample M identical to pole B (PB). Samples G and G* were replicated samples evaluated by different groups of consumers.

1 observed (Fig. 4). The position of sample J differed between
2 sample configurations. It was located in a distinct position
3 in the evaluation of the whole set (Fig. 4A) but was located
4 near sample L when two consumer groups evaluated the
5 split set (Fig. 4B). Besides, sample M was relatively close to
6 sample I when considering whole sample set evaluation but
7 it was in a distinct position when the two consumer groups
8 evaluated the split set. As shown in Fig. 4B, the projection of
9 sample G* (considered as supplementary individual on the
10 sensory space) was located relatively far from sample G.
11 Besides, conclusions regarding the degree of similarity
12 between samples and poles were not affected by the type of
13 evaluation (Fig. 4).

14 Both PSP approaches were able to spot samples identical
15 to the poles, regardless of the type of evaluation (whole
16 sample set and data aggregation from the evaluation of the
17 split sets). As shown in Fig. 1, samples J and M were located
18 opposite to the direction of increasing difference with poles
19 PA and PB, respectively, when PSP was considered (Fig. 3).
20 Meanwhile, in t-PSP samples J and M were located close to
21 the columns that represent similarity to poles A and B (PA+
22 and PB+), respectively, in both types of evaluations (Fig. 4).

23 The agreement between sample configurations from PSP
24 with scales and t-PSP was not good. The RV coefficient
25 between sample configurations was low and non-significant
26 (Table 4).

28 DISCUSSION

29 The main advantage of PSP over other holistic methodolo-
30 gies for sensory characterization is the fact that it could
31 potentially allow data aggregation from different sessions,
32 because of the comparative nature of the task with fixed refer-
33 ences and the sequential monadic presentation of the
34 tested samples. Nevertheless, this issue has not been previ-
35 ously explored in the literature. In this context, the present
36 research compared sample configurations obtained through
37 the evaluation of the whole set and data aggregation from
38 the evaluation of the split set by different consumer groups
39 using two PSP approaches: PSP with scales and t-PSP.

40 The RV coefficients between sample configurations from
41 the evaluation of the whole set and aggregated data from
42 the evaluation of the split set by different consumer groups
43 were significant and higher than 0.79 (Table 4). These RV
44 coefficients can be regarded as indicator of good agreement
45 between sample configurations (Faye *et al.* 2004; Abdi *et al.*
46 2007; Lelièvre *et al.* 2008). Moreover, most of the conclu-
47 sions regarding the degree of similarity between samples
48 and the poles did not change with the type of evaluation
49 nor did the ability of the methodology to spot samples
50 identical to the poles when presented blind to consumers
51 (c.f. Figs. 1–4). Therefore, it can be inferred that aggregation

52 of data collected in different sessions with different con-
53 sumers provides similar information than the evaluation of
54 the whole sample set.

55 Despite the relatively high RV coefficients, it is important
56 to highlight that differences existed in some of the conclu-
57 sions from the sensory characterizations obtained from the
58 evaluation of the whole sample set and data aggregation
59 from the evaluation of the split set. Firstly, differences in the
60 relative position of the samples were identified between
61 sample configurations obtained through the evaluation of
62 the whole set and aggregated data from the evaluation of
63 the split set, which led to different conclusions regarding
64 similarities and differences among samples. When PSP with
65 scales was used in study 2 differences in conclusions regard-
66 ing similarities and differences among samples affected a
67 larger proportion of the samples (Fig. 3). Furthermore, dif-
68 ference scores between samples and the poles were affected
69 by the type of evaluation in one of the studies. As shown in
70 Table 5, the way in which consumers used the scale to rate
71 the degree of difference between samples and the poles in
72 study 1 significantly differed depending on whether they
73 evaluated the whole or the split sample set.

74 It is important to note that the influence of data aggrega-
75 tion was larger in study 2 than in study 1, which could be
76 related to the fact that differences among samples were
77 smaller in the former study. It could be hypothesized that
78 the influence of data aggregation grows as the degree of dif-
79 ference among sample decreases. For this reason, studies
80 aimed at investigating the influence of the degree of differ-
81 ence among samples on data aggregation from PSP seem
82 necessary in order to make recommendations on the appli-
83 cability of the methodology.

84 In study 2 the sample evaluated by the two consumer
85 groups was located far from each other on the sensory map
86 obtained from data aggregation when samples were evalu-
87 ated using t-PSP but close together when PSP with scales
88 was considered (c.f. Figs. 3B and 4B). Conclusions regarding
89 the sensory characteristics of the replicated sample differed
90 between consumer groups. The previous results raise con-
91 cerns about data aggregation from t-PSP and indicate that
92 further research should be carried out.

93 In the present study, sample configurations obtained
94 using PSP with scales and t-PSP were largely similar in
95 study 1 in agreement with results reported by Ares *et al.*
96 (2013). However, sample configurations obtained using PSP
97 with scales and t-PSP differed and were not significantly
98 correlated in study 2 (c.f. Figs. 3 and 4). This can be attrib-
99 uted to the fact that sensory differences among samples
100 were smaller in study 2 than in study 1. Differences could be
101 also related to the sensory complexity of products. In this
102 sense, study 2 involved more complex products than study
103 1. In the present study it is not possible to establish which of
104 these methodologies provided more valid information

regarding similarities and differences among samples due to the fact that commercial samples were evaluated. Further research comparing sample configurations from PSP with scales and t-PSP with those obtained using descriptive analysis with trained assessors would be necessary to determine the validity of sensory characterizations obtained using PSP approaches.

Ares *et al.* (2013) reported that consumers found a modified version of t-PSP easier than PSP with scales. t-PSP can be considered as a more intuitive methodology than PSP with scales. Although PSP with scales require assessors to use unstructured scales to evaluate the degree of difference between samples and each of the poles, t-PSP only requires assessors to indicate to which of the poles each sample resembles the most and to which it resembles the least. This must not be the case if working with trained assessors that are accustomed to scaling since for them the use of scales can be perceived as easier than t-PSP. Differences between trained and untrained assessors in the performance of PSP approaches are worth studying.

Regarding the influence of the PSP approach on the quality of data aggregation, t-PSP provided better results than PSP with scales when sample configurations were considered. As shown in Figs. 1–4, in both studies sample configurations obtained by data aggregation from the evaluation of the split set were more similar to sample configurations from the evaluation of the whole set when t-PSP was considered than for PSP with scales. On the contrary, the opposite trend was found when the position of a replicated sample was considered in study 2. This difference can be attributed to the nature of the data of both approaches. PSP with scales rely on average data from continuous scales, whereas t-PSP is based on frequencies.

CONCLUSIONS

Results from the present work confirmed that data aggregation of data collected in different sessions with different assessors using PSP provides similar information to the evaluation of samples in a single session. This characteristic of PSP makes it a particularly interesting alternative when using consumer-based characterizations for new product development. However, it is important to highlight that some differences in the conclusions obtained from the evaluation of the whole and split set were identified, particularly when PSP with scales was used to evaluate samples in study 2.

These results suggest that data aggregation seems feasible when working with samples that are markedly different and that further research would be needed to confirm and extend the findings of the present work. In this sense, it would be interesting to investigate how product complexity, number of samples in the set and the degree of differences

among samples affect the quality of sample configurations obtained by aggregating data from the evaluation of split sample sets by different consumer groups. Furthermore, research comparing the validity, reproducibility and data aggregation of PSP with scales and t-PSP in different product categories of different complexity is still necessary.


Regarding the comparison of PSP approaches in the present study better agreement between sample configurations was obtained when t-PSP was used, which is probably related to the fact that this methodology does not deal with the heterogeneity in consumers' use of the scale (Ares *et al.* 2011).

ACKNOWLEDGMENTS

The authors are indebted to Comisión Sectorial de Investigación Científica (Universidad de la República, Uruguay) for financial support and for the scholarship granted to author L.A. to perform a short stay at Instituto de Agroquímica y Tecnología de Alimentos (Valencia, Spain).

REFERENCES

- ABDI, H., VALENTIN, D., CHOLLET, S. and CHREA, C. 2007. Analyzing assessors and products in sorting studies: DISTATIS, theory and applications. *Food Qual. Prefer.* 18, 627–640.
- ARES, G. and VARELA, P. 2014. Comparison of novel methodologies for sensory characterization. In *Novel Techniques in Sensory Characterization and Consumer Profiling* (P. Varela and G. Ares, eds.) pp. 365–389, CRC Press, Boca Raton, FL.
- ARES, G., GIMÉNEZ, A. and BRUZZONE, F. 2011. Is a consumer panel able to reliably evaluate the texture of dairy desserts using unstructured intensity scales? Evaluation of global and individual performance. *J. Sensory Studies* 26, 363–370.
- ARES, G., de SALDAMANDO, L., VIDAL, L., ANTÚNEZ, L., GIMÉNEZ, A. and VARELA, P. 2013. Polarized projective mapping: Comparison with Polarized Sensory Positioning approaches. *Food Qual. Prefer.* 28, 510–518.
- CHOLLET, S. and VALENTIN, D. 2001. Impact of training on beer flavour perception and description: Are trained and untrained subjects really different. *J. Sensory Studies* 16, 601–618.
- CHOLLET, S., LELIÈVRE, M., ABDI, H. and VALENTIN, D. 2011. Sort and beer: Everything you wanted to know about the sorting task but did not dare to ask. *Food Qual. Prefer.* 22, 507–520.
- de SALDAMANDO, L., DELGADO, J., HERENCIA, P., GIMÉNEZ, A. and ARES, G. 2013. Polarized sensory positioning: Do conclusions depend on the poles? *Food Qual. Prefer.* 29, 25–32.







| | | | |
|----|---|--|----|
| 1 | DEHLHOLM, C., BROCKHOFF, P.B., MEJNERT, L., AASLYNG, | MURRAY, J.M., DELAHUNTY, C.M. and BAXTER, I.A. 2001. | 32 |
| 2 | M.D. and BREDIE, W.L.P. 2012. Rapid descriptive sensory | Descriptive sensory analysis: Past, present and future. <i>Food</i> | 33 |
| 3 | methods – comparison of free multiple sorting, partial | <i>Res. Int.</i> 34, 461–471. | 34 |
| 4 | napping, napping, flash profiling and conventional profiling. | R CORE TEAM. 2013. <i>R: A Language and Environment for</i> | 35 |
| 5 | <i>Food Qual. Prefer.</i> 26, 267–277. | <i>Statistical Computing</i> , R Foundation for Statistical | 36 |
| 6 | FAYE, P., BRÉMAUD, D., DURAND-DAUBIN, D., | Computing, Vienna. | 37 |
| 7 | COURCOUX, P., GIBOREAU, A. and NICOD, A. 2004. | RISVIK, E., MCEWAN, J.A. and RODBOTTEN, M. 1997. | 38 |
| 8 | Perceptive free sorting and verbalization tasks with naive | Evaluation of sensory profiling and projective mapping | 39 |
| 9 | subjects: An alternative to descriptive mappings. <i>Food Qual.</i> | <i>data. Food Qual. Prefer.</i> 8, 63–71. | 40 |
| 10 |  <i>Prefer.</i> 15, 781–791. | ROBERT, P. and ESCOUFIER, Y. 1976. A unifying tool for linear | 41 |
| 11 | ISO. 2007. <i>Sensory analysis: General guidance for the design of test</i> | multivariate statistical methods: The RV coefficient. <i>Appl.</i> | 42 |
| 12 | <i>rooms, ISO Standard 8589</i> . International Organization for | <i>Stat.</i> 25, 257–265. | 43 |
| 13 | Standardization, Geneva. | STONE, H., SIDEL, J.L., OLIVER, S., WOOLSEY, A. and | 44 |
| 14 | JOSSE, J., PAGÉS, J. and HUSSON, F. 2008. Testing the | SINGLETON, R.C. 1974. Sensory evaluation by quantitative | 45 |
| 15 | significance of the RV coefficient. <i>Comput. Stat. Data Anal.</i> | descriptive analysis. <i>Food Technol.</i> 28, 24–33. | 46 |
| 16 | 53, 82–91. | TEILLET, E. 2014. Polarized Sensory Positioning methodologies. | 47 |
| 17 | LAWLESS, H.T. and HEYMANN, H. 2010. <i>Sensory Evaluation of</i> | <i>In Novel Techniques in Sensory Characterization and Consumer</i> | 48 |
| 18 | <i>Food. Principles and Practices</i> , 2nd Ed., Springer, New York, | <i>Profiling</i> (P. Varela and G. Ares, eds.) pp. 255–269, CRC Press, | 49 |
| 19 | NY. | Boca Raton, FL. | 50 |
| 20 | LELIÈVRE, M., CHOLLET, S., ABDI, H. and VALENTIN, D. | TEILLET, E., SCHLICH, P., URBANO, C., CORDELLE, S. and | 51 |
| 21 | 2008. What is the validity of the sorting study for describing | GUICHARD, E. 2010. Sensory methodologies and the taste of | 52 |
| 22 | beers? A study using trained and untrained assessors. <i>Food</i> | water. <i>Food Qual. Prefer.</i> 21, 967–976. | 53 |
| 23 | <i>Qual. Prefer.</i> 19, 697–703. | VALENTIN, D., CHOLLET, S., LELIEVRE, M. and ABDI, H. | 54 |
| 24 | LÊ, S., JOSSE, J. and HUSSON, F. 2008. FactoMineR: An R | 2012. Quick and dirty but still pretty good: A review of new | 55 |
| 25 | package for multivariate analysis. <i>J. Stat. Softw.</i> 25, 1–18. | descriptive methods in food science. <i>Int. J. Food Sci. Technol.</i> | 56 |
| 26 | MACRAE, A.W., HOWGATE, P. and GEELHOED, E.N. 1990. | 47, 1563–1578. | 57 |
| 27 | Assessing the similarity of odours by sorting and by triadic | VARELA, P. and ARES, G. 2012. Sensory profiling, the blurred | 58 |
| 28 | comparison. <i>Chem. Senses</i> 15, 661–699. | line between sensory and consumer science. A review of novel | 59 |
| 29 | MEILGAARD, M.C., CIVILLE, G.V. and CARR, B.T. 1999. | methods for product characterization. <i>Food Res. Int.</i> 48, | 60 |
| 30 | <i>Sensory Evaluation Techniques</i> , 2nd Ed., CRC Press, Boca | 893–908. | 61 |
| 31 | Raton, FL. | | |

AUTHOR QUERY FORM

Dear Author,

During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof.

Many thanks for your assistance.

| Query References | Query | Remarks |
|------------------|--|---|
| 1 | AUTHOR: Please confirm that given names (red) and surnames/family names (green) have been identified correctly. |  |
| 2 | AUTHOR: Please supply postcode for affiliation 1. |  |
| 3 | AUTHOR: Risvik <i>et al.</i> 1994 has been changed to Risvik <i>et al.</i> 1997 so that this citation matches the Reference List. Please confirm that this is correct. |  |
| 4 | AUTHOR: Hopfer and Heymann 2013 has not been included in the Reference List. Please supply full publication details. |  |
| 5 | AUTHOR: Teillet <i>et al.</i> 2014 has not been included in the Reference List. Please supply full publication details. |  |
| 6 | AUTHOR: Please check if the first name initials in Reference Chollet <i>et al.</i> , 2011 for author Lelièvre are correct. |  |