Journal of Sensory Studies

Journal of Sensory Studies ISSN 0887-8250

EVALUATION OF DATA AGGREGATION IN POLARIZED SENSORY POSITIONING

LUCÍA ANTÚNEZ[1,4], ANA SALVADOR[2], LUIS de SALDAMANDO[1], PAULA VARELA[3], ANA GIMÉNEZ[1] and GASTÓN ARES[1]

[1]Departamento de Ciencia y Tecnología de Alimentos, Facultad de Química, Universidad de la República, Montevideo CP 11800, Uruguay
[2]Instituto de Agroquímica y Tecnología de Alimentos (CSIC), Paterna, Valencia, Spain
[3]Nofima AS, Ås, Norway

[4]Corresponding 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 less sophisticated and faster methodologies has markedly grown in the last decade (Valentin *et al.* 2012; Varela and Ares 2012).

Holistic methodologies, such as sorting and projective mapping, are one of the novel approaches for sensory characterization (Valentin *et al.* 2012). They are based on the evaluation of global differences among samples, enabling identification of the main sensory characteristics responsible for perceived similarity among samples (Ares and Varela 2014). Despite the fact that these methodologies have been reported to provide valid and reliable information (Risvik *et al.* 1997; Chollet and Valentin 2001; Faye *et al.* 2004; Chollet *et al.* 2011; Dehlholm *et al.* 2012; Hopfer and Heymann 2013), one of their main disadvantages is that the entire set of products must be simultaneously evaluated in the same session (Teillet *et al.* 2010; Varela and Ares 2012; Ares *et al.* 2013). This restricts the number of samples that can be evaluated as well as the possibility of comparing samples evaluated in different moments in time.

In order to overcome this limitation, Teillet *et al.* (2010) have developed polarized sensory positioning (PSP). This methodology is based on the comparison of samples with a fixed set of reference products called “poles” (Teillet 2014). Despite the fact that this methodology was originally developed to explore the sensory characteristic of water with trained assessors (Teillet *et al.* 2010), it has been successfully used with naïve consumers (Ares *et al.* 2013; de Saldamando *et al.* 2013; Teillet 2014).

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

Regardless of the specific approach considered for sample evaluation, PSP approaches are based on the comparison of samples with a set of poles, which are kept constant across different sessions. Therefore, PSP makes it possible to aggregate data from different sessions and to accumulate data from different studies. Considering the increasing interest in consumer-based sensory characterizations and the difficulties usually encountered for recruiting consumers for replicated evaluations, the possibility of aggregating data from sensory characterization studies performed with different consumers is particularly interesting in both industrial and academic applications. However, to the authors’ knowledge no study evaluating data aggregation from sensory characterization with consumers using PSP has been published in refereed journals.

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

MATERIALS AND METHODS

Two studies with two different product categories (orange-flavored powdered drinks and chocolate milk beverages) with 240 consumers were conducted to evaluate data aggregation in PSP. In each study two PSP approaches were considered: PSP with scales and t-PSP. For each PSP approach a between-subject design was used to compare sample configurations for the evaluation of the whole and the split sample sets.

Samples

A total of seven samples of commercial orange-flavored powdered drinks were used in study 1. All samples were available in the Uruguayan market and were purchased from local supermarkets in Montevideo (Uruguay). The set involved six samples (A–F) and a set of three poles (PA, PB and PC). Two of the poles were identical to the samples in order to evaluate the validity of the methodology. Poles were selected based on results from a previous study that used the projective mapping methodology to identify the sensory characteristics responsible for the main differences among commercial samples of orange-flavored powdered drinks (Ares *et al.* 2013). Three main groups of samples were identified in that study: one was characterized by its low total flavor intensity, a second one by its sourness and a third group was described as sweet and with intense orange flavor. Considering these results one pole was selected from each of those groups: PB was characterized by its low total flavor intensity, PC was described as a sweet drink with intense orange flavor and PA was characterized by its sourness (Ares *et al.* 2013). Table 1 provides a description of the samples in terms of their main characteristics and market positioning. Samples were prepared following the recommendations provided by the manufacturer on the package. The powders were diluted in tap water and stored in a fridge at 10C until they were served to consumers within 4 h.

Study 2 was carried out with samples of commercial chocolate milk beverages available in the Spanish market. The set involved a total of seven samples (G–M). Three poles (A, B and C) were selected based on results from a

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.

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

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.

Data Collection

In each study consumers were randomly divided into two groups of 120, each of which 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

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

Data Analysis

PSP with Scales. Data from PSP with scales were considered as sensory descriptors and consequently analyzed using principal component analysis (PCA) (Teillet 2014). For each sample, the average score was calculated and a matrix containing samples in rows and poles in columns was constructed. PCA was applied on the correlation matrix of average scores. When samples were evaluated by different groups of consumers, data were analyzed by binding the matrices obtained for each consumer group.

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

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

t-PSP. Data from t-PSP were analyzed considering the pole to which the sample resemble the most (named A+, B+ or C+) and the pole to which the sample resemble the least (named A–, B– and C–) as qualitative variables. A frequency table containing the number of times a sample was regarded as most similar and most different to each of the poles was constructed and analyzed by means of correspondence analysis (CA) (Teillet *et al.* 2014). When samples were evaluated by different groups of consumers, data were analyzed by binding the frequency tables obtained for each consumer group.

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

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

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

RESULTS

Study 1

As shown in Figs. 1 and 2, regardless of the PSP approach used for evaluating samples, the percentage of variance explained by the first and second 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 methodologies, sample configurations obtained through data aggregation from the evaluation of the split set by different consumer groups was similar to those obtained from the evaluation of the whole set. The RV coefficients between sample configurations were significant and higher than 0.90 (Table 4). Besides, for both PSP with scales and t-PSP the position of the samples with respect to the poles was similar when the configuration was based on data from the evaluation of the whole set or based on data aggregation from the evaluation of the split set (Figs. 1 and 2, respectively).

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

Despite the high similarity between sample configurations, some differences in conclusions regarding similarities and differences among samples were identified. For PSP with scales, the relative position of sample D in the sensory space markedly differed between sample configurations obtained through the evaluation of the whole set and the data aggregation from the evaluation of the split set by different consumer groups. When the whole set was evaluated, sample D was located in a distinct position (Fig. 1A), while when sample configurations obtained by aggregating data from the evaluation of the split set was considered, sample D was regarded as similar to samples E and C (Fig. 1B).

ANOVA was used to assess if dissimilarity scores obtained by the evaluation of the whole sample set and data aggregation from the evaluation of the split set by different consumer groups significantly differed. As shown in Table 5, difference scores between samples and poles A and C were significantly affected by the type of evaluation (whole sample set or data aggregation from the evaluation of the split set). On average, difference scores were higher when consumers evaluated the whole sample set than when

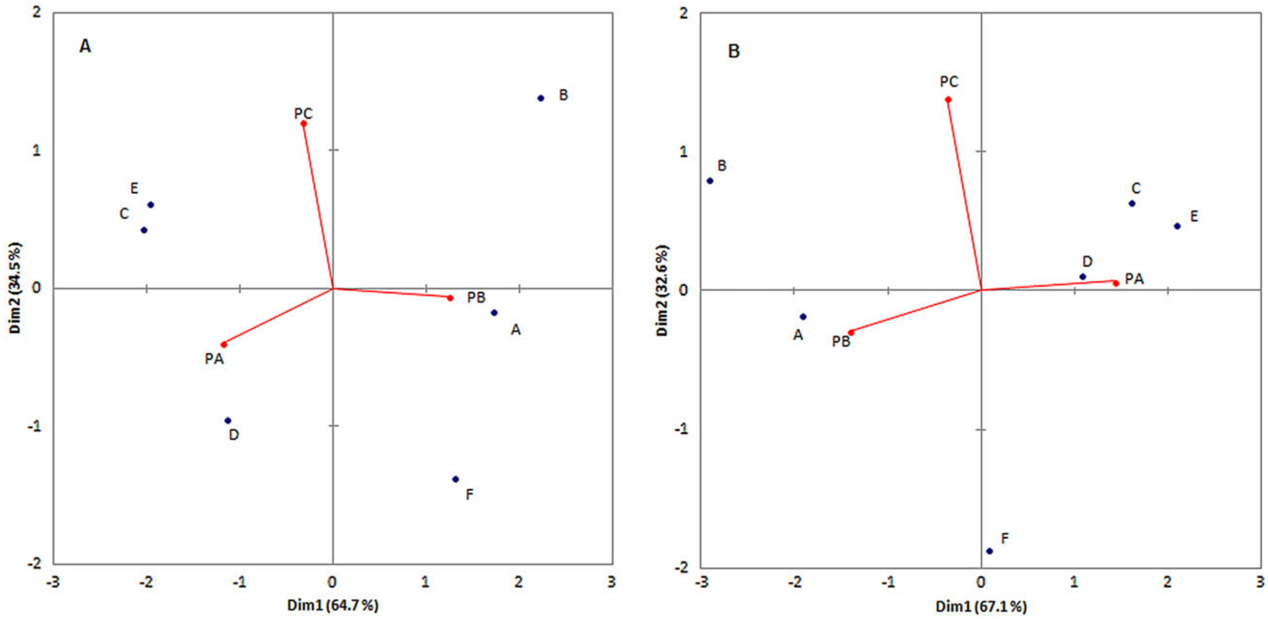

FIG. 1. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF PRINCIPAL COMPONENT ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING POLARIZED SENSORY POSITIONING WITH SCALES FOR THE EVALUATION OF ORANGE-FLAVORED 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

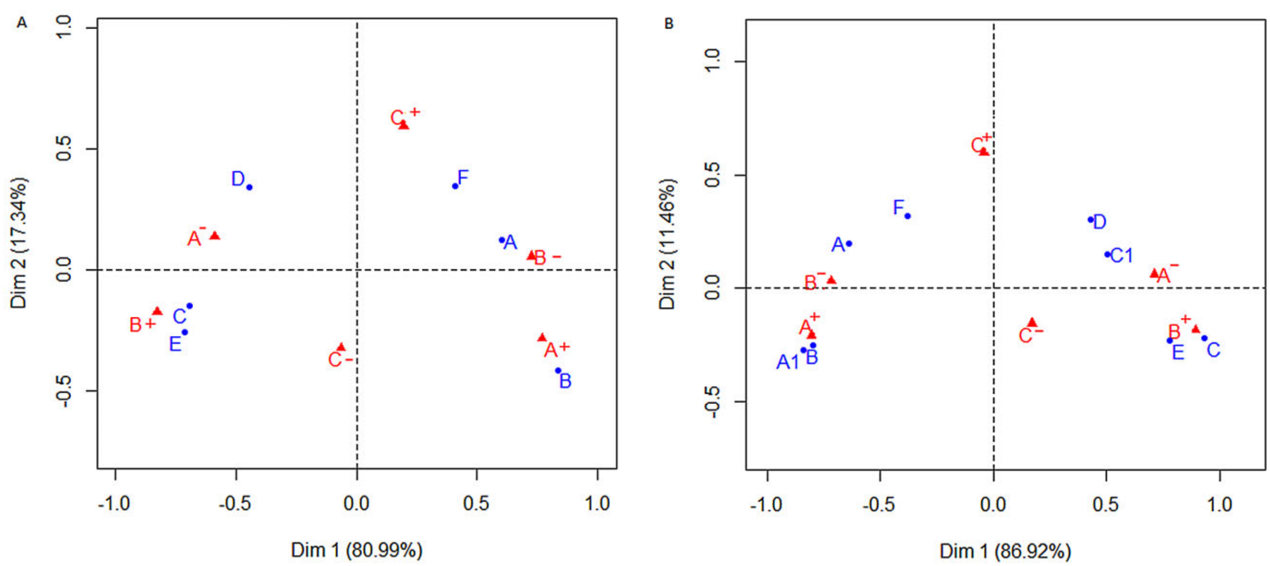

FIG. 2. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF CORRESPONDENCE ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING TRIADIC POLARIZED SENSORY POSITIONING FOR THE EVALUATION OF ORANGE-FLAVORED POWDERED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

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

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
	Whole set	t-PSP versus PSP with scales	0.96	0.010
Study 2	Whole set versus data aggregation from the split set	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.

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).

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

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

		Pole		
Study	Source of variation	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 ($P = 0.1470$)

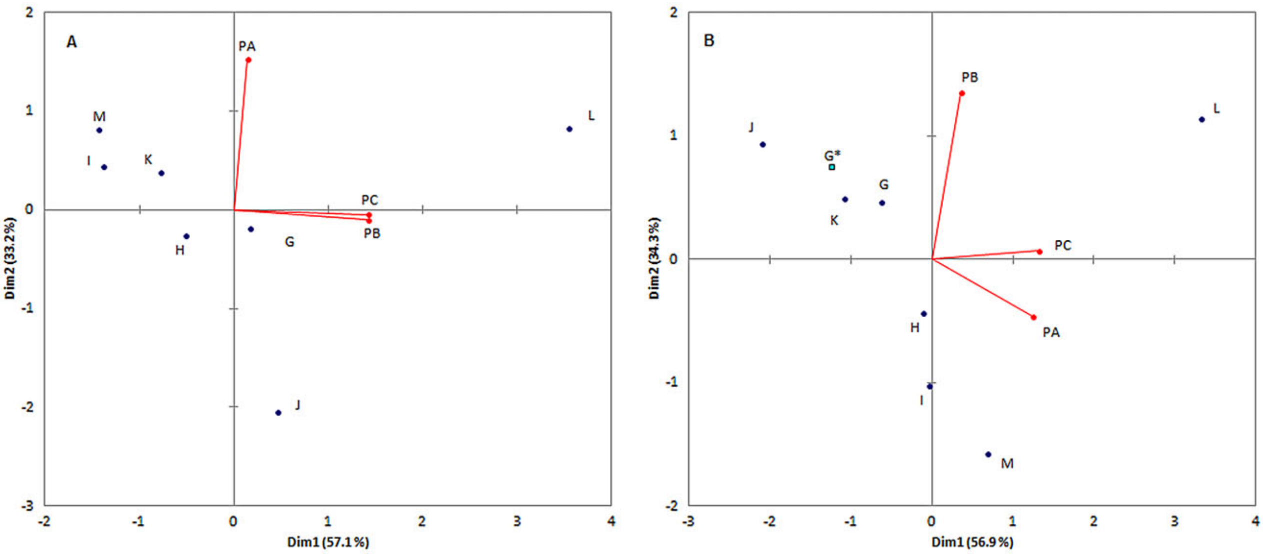

FIG. 3. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF PRINCIPAL COMPONENT ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING POLARIZED SENSORY POSITIONING WITH SCALES FOR THE EVALUATION OF CHOCOLATE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

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 observed (Fig. 4). The position of sample J differed between

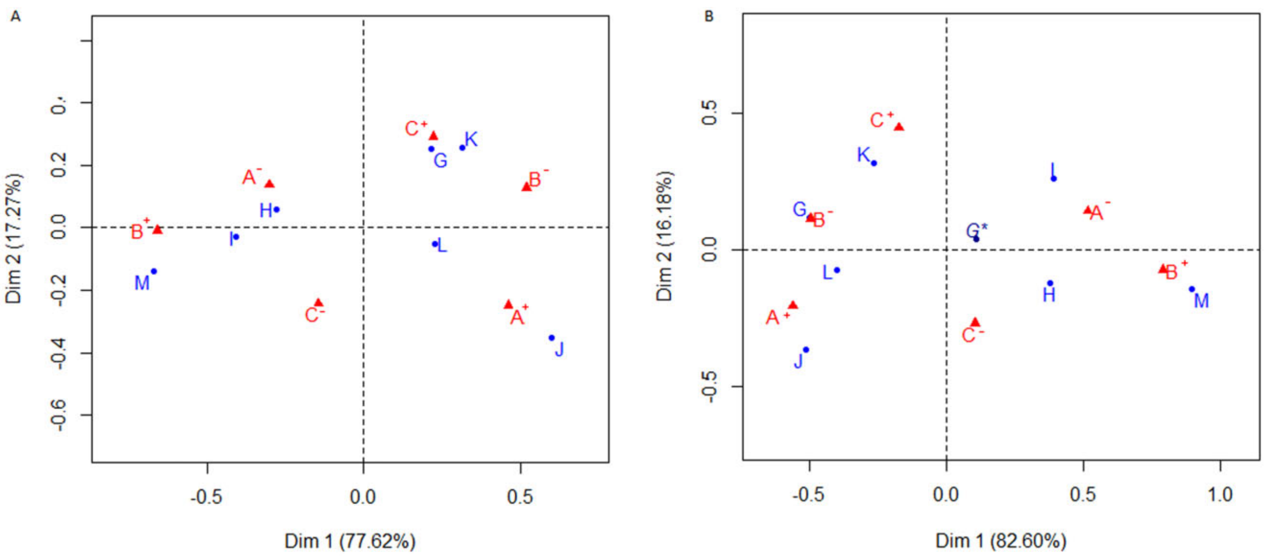

FIG. 4. SAMPLE CONFIGURATIONS ON THE FIRST AND SECOND DIMENSIONS OF CORRESPONDENCE ANALYSIS OBTAINED THROUGH THE EVALUATION OF THE WHOLE SET (A) AND DATA AGGREGATION FROM THE SPLIT SET BY DIFFERENT CONSUMER GROUPS (B) USING TRIADIC POLARIZED SENSORY POSITIONING FOR THE EVALUATION OF CHOCOLATE MILK BEVERAGES
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.

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

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

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

DISCUSSION

The main advantage of PSP over other holistic methodologies for sensory characterization is the fact that it could potentially allow data aggregation from different sessions, because of the comparative nature of the task with fixed references and the sequential monadic presentation of the tested samples. Nevertheless, this issue has not been previously explored in the literature. In this context, the present research compared 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 two PSP approaches: PSP with scales and t-PSP.

The RV coefficients between sample configurations from the evaluation of the whole set and aggregated data from the evaluation of the split set by different consumer groups were significant and higher than 0.79 (Table 4). These RV coefficients can be regarded as indicator of good agreement between sample configurations (Faye *et al.* 2004; Abdi *et al.* 2007; Lelièvre *et al.* 2008). Moreover, most of the conclusions regarding the degree of similarity between samples and the poles did not change with the type of evaluation nor did the ability of the methodology to spot samples identical to the poles when presented blind to consumers (c.f. Figs. 1–4). Therefore, it can be inferred that aggregation of data collected in different sessions with different consumers provides similar information than the evaluation of the whole sample set.

Despite the relatively high RV coefficients, it is important to highlight that differences existed in some of the conclusions from the sensory characterizations obtained from the evaluation of the whole sample set and data aggregation from the evaluation of the split set. Firstly, differences in the relative position of the samples were identified between sample configurations obtained through the evaluation of the whole set and aggregated data from the evaluation of the split set, which led to different conclusions regarding similarities and differences among samples. When PSP with scales was used in study 2 differences in conclusions regarding similarities and differences among samples affected a larger proportion of the samples (Fig. 3). Furthermore, difference scores between samples and the poles were affected by the type of evaluation in one of the studies. As shown in Table 5, the way in which consumers used the scale to rate the degree of difference between samples and the poles in study 1 significantly differed depending on whether they evaluated the whole or the split sample set.

It is important to note that the influence of data aggregation was larger in study 2 than in study 1, which could be related to the fact that differences among samples were smaller in the former study. It could be hypothesized that the influence of data aggregation grows as the degree of difference among sample decreases. For this reason, studies aimed at investigating the influence of the degree of difference among samples on data aggregation from PSP seem necessary in order to make recommendations on the applicability of the methodology.

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

In the present study, sample configurations obtained using PSP with scales and t-PSP were largely similar in study 1 in agreement with results reported by Ares *et al.* (2013). However, sample configurations obtained using PSP with scales and t-PSP differed and were not significantly correlated in study 2 (c.f. Figs. 3 and 4). This can be attributed to the fact that sensory differences among samples were smaller in study 2 than in study 1. Differences could be also related to the sensory complexity of products. In this sense, study 2 involved more complex products than study 1. In the present study it is not possible to establish which of 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.

DEHLHOLM, C., BROCKHOFF, P.B., MEJNERT, L., AASLYNG, M.D. and BREDIE, W.L.P. 2012. Rapid descriptive sensory methods – comparison of free multiple sorting, partial napping, napping, flash profiling and conventional profiling. Food Qual. Prefer. *26*, 267–277.

FAYE, P., BRÉMAUD, D., DURAND-DAUBIN, D., COURCOUX, P., GIBOREAU, A. and NICOD, A. 2004. Perceptive free sorting and verbalization tasks with naive subjects: An alternative to descriptive mappings. Food Qual. Prefer. *15*, 781–791.

HOPFER, H. and HEYMANN, H. 2013. A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. Food Qual. Prefer. *28*, 164–181.

ISO. 2007. *Sensory analysis: General guidance for the design of test rooms, ISO Standard 8589.* International Organization for Standardization, Geneve.

JOSSE, J., PAGÉS, J. and HUSSON, F. 2008. Testing the significance of the RV coefficient. Comput. Stat. Data Anal. *53*, 82–91.

LAWLESS, H.T. and HEYMANN, H. 2010. *Sensory Evaluation of Food. Principles and Practices*, 2nd Ed., Springer, New York, NY.

LELIÈVRE, M., CHOLLET, S., ABDI, H. and VALENTIN, D. 2008. What is the validity of the sorting study for describing beers? A study using trained and untrained assessors. Food Qual. Prefer. *19*, 697–703.

LÊ, S., JOSSE, J. and HUSSON, F. 2008. FactoMineR: An R package for multivariate analysis. J. Stat. Softw. *25*, 1–18.

MACRAE, A.W., HOWGATE, P. and GEELHOED, E.N. 1990. Assessing the similarity of odours by sorting and by triadic comparison. Chem. Senses *15*, 661–699.

MEILGAARD, M.C., CIVILLE, G.V. and CARR, B.T. 1999. *Sensory Evaluation Techniques*, 2nd Ed., CRC Press, Boca Raton, FL.

MURRAY, J.M., DELAHUNTY, C.M. and BAXTER, I.A. 2001. Descriptive sensory analysis: Past, present and future. Food Res. Int. *34*, 461–471.

R CORE TEAM. 2013. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna.

RISVIK, E., MCEWAN, J.A. and RODBOTTEN, M. 1997. Evaluation of sensory profiling and projective mapping data. Food Qual. Prefer. *8*, 63–71.

ROBERT, P. and ESCOUFIER, Y. 1976. A unifying tool for linear multivariate statistical methods: The RV coefficient. Appl. Stat. *25*, 257–265.

STONE, H., SIDEL, J.L., OLIVER, S., WOOLSEY, A. and SINGLETON, R.C. 1974. Sensory evaluation by quantitative descriptive analysis. Food Technol. *28*, 24–33.

TEILLET, E. 2014. Polarized Sensory Positioning methodologies. In *Novel Techniques in Sensory Characterization and Consumer Profiling* (P. Varela and G. Ares, eds.) pp. 255–269, CRC Press, Boca Raton, FL.

TEILLET, E., SCHLICH, P., URBANO, C., CORDELLE, S. and GUICHARD, E. 2010. Sensory methodologies and the taste of water. Food Qual. Prefer. *21*, 967–976.

VALENTIN, D., CHOLLET, S., LELIEVRE, M. and ABDI, H. 2012. Quick and dirty but still pretty good: A review of new descriptive methods in food science. Int. J. Food Sci. Technol. *47*, 1563–1578.

VARELA, P. and ARES, G. 2012. Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. Food Res. Int. *48*, 893–908.