Comparison of rapid sensory characterization methodologies for the development of functional yogurts

Rafael S. Cadena a,⁎, Daniel Caimí a, Isabel Jaunarena a, Inés Lorenzo a, Leticia Vidal a, Gastón Aresa a, Rosires Deliza b, Ana Giménez a

a Departamento de Ciencia y Tecnología de Alimentos, Facultad de Química, Universidad de la República, Gral. Flores 2124, C.P. 11800, Montevideo, Uruguay
b Embrapa Food Technology, Av. das Américas 29501, CEP 23020-470, Rio de Janeiro, RJ, Brazil

1. Introduction

Probiotics and prebiotics are one of the most profitable categories within the functional food market (Bigiardi & Galati, 2013; Cruz et al., 2010). Combinations of these functional ingredients are increasingly incorporated into food products, fermented milk being the most popular vehicle (Al-Sheraji et al., 2013). Probiotics are live microorganisms, which confer health benefits to the host when consumed in adequate quantities (Guarner & Schaffandi, 1998). Prebiotics are short-chain carbohydrates non-digestible in the human gastrointestinal tract that enhance the activity of intestinal flora and exert health benefits to the health (Quigley, Hudson, & Englyst, 1999). The addition of probiotics and prebiotics to food products can modify their sensory characteristics which can decrease consumer overall liking (Cruz et al., 2010; Gallardo-Escamilla, Kelly, & Delahunty, 2005; La Torre, Tamime, & Muir, 2003; Luckow, Sheehan, Fitzgerald, & Delahunty, 2006).

Consumers should ingest functional foods on a regular basis to achieve the health benefits derived from them, so the sensory characteristics of functional foods must not discourage sustained consumption (Sarubin, 2000; Siró, Kápolna, Kápolna, & Lugasi, 2008). Therefore, food companies should rely on valid methodologies to assess the impact of functional ingredients on the sensory characteristics of the products.

The sensory characteristics of products have been traditionally assessed using descriptive analysis with trained assessors (Lawless & Heymann, 2010; Stone & Sidel, 2004). Although this methodology provides detailed, reliable and reproducible results, it is time consuming (Murray, Delahunty, & Baxter, 2001).

Functional food development is a complex, expensive and risky process, which involves long-term studies to gather scientific evidence of their health effects (Jones & Jew, 2007). Therefore, it is important to speed up the product development process and to assure the success of the developed products from the early stages of product development. In this context, methodologies that allow gathering information about the sensory characteristics of products in short time frames, and directly from consumers, are valuable tools.

Several rapid methodologies for sensory characterization have been recently developed. These methodologies can be performed without prior training, which makes them simple and flexible alternatives for sensory characterization with both trained assessors and consumers (Varela & Ares, 2012). They can be divided into three main types: methodologies based on the evaluation of specific attributes, holistic...
methodologies, based on global similarities and differences between products, and methodologies based on the comparison of products with references (Valentin, Chollet, Leleivre, & Abdi, 2012).

A check-all-that-apply (CATA) question is one of the most novel methodologies based on the evaluation of specific attributes (Adams, Williams, Lancaster, & Foley, 2007). A CATA question consists of a list of words or phrases from which respondents should select all that they consider appropriate to describe a product (Varela & Ares, 2012). This methodology has been reported to be a simple, valid and reproducible alternative for gathering information about the sensory characteristics of a wide range of products (Bruzzone, Ares, & Giménez, 2012; Dooley, Lee, & Meullenet, 2010; Jaeger et al., 2013; Meyners, Castura, & Carr, 2013; Parente, Manzoni, & Ares, 2011; Plaehn, 2012).

Projective mapping (PM) or Napping® is a holistic method based on assessors’ individual perception of overall similarities and dissimilarities among products. Assessors are asked to provide a two dimensional representation of a group of samples, according to their own criteria (Risvik, McEvan, Colwill, Rogers, & Lyon, 1994). In this representation, the Euclidean distance between each pair of samples is a measure of their dissimilarity. The criteria used by assessors to locate samples depend on the relative importance they attach to their sensory characteristics, which makes projective mapping a flexible and spontaneous methodology (Varela & Ares, 2012).

Polarized sensory positioning (PSP) is a reference-based method that has been developed by Teillet, Schlich, Urbano, Cordelle, and Guichard (2010) for sensory characterization of mineral water. The methodology is based on the evaluation of the global difference between samples and a fixed set of reference products, named poles (Teillet, 2014). The main advantage of this methodology is the possibility of aggregating data collected in different sessions (Ares & Varela, 2014).

Compared to descriptive analysis, rapid methodologies for sensory characterization have been used for a relatively short period of time and have been used in a limited number of applications. Therefore, research on their applicability, reliability, and reproducibility for sensory characterization of products with different sensory complexity is still needed in order to allow them to be established as standard tools in sensory and consumer science (Ares & Varela, 2014).

In this context, the aim of the present work was to compare three rapid methodologies for sensory characterization (check-all-that-apply questions, projective mapping and polarized sensory positioning) with descriptive analysis during the development of low-fat functional yogurts, enriched with probiotics and prebiotics.

### 2. Material and methods

#### 2.1. Samples

Eight low-fat yogurts enriched with a prebiotic ingredient were formulated following a 2\(^3\) full factorial design with the following factors: sugar concentration (4.0% vs. 8.0%), prebiotic ingredient (native inulin — Frutaft IQ, Sensus, Netherlands, and fructooligosaccharide (FOS) — Orafti® P95, Beneo GmbH, Mannheim, Germany) and stabilizer concentration (Dairy Blend YG LP, TIC Gums, White Marsh, Maryland, USA). The concentration of the prebiotic ingredient in all formulations was 6.0%. A similar percentage of inulin has been considered in the development of milk desserts (Tárrega, Rocafull, & Costell, 2010). Besides, all yogurts contained 1% modified starch (National 465, National Starch, Trombudo Central, Santa Catarina, Brazil) and 2% skim milk powder (Conaprole, Montevideo, Uruguay). The rest of the formulation consisted of skimmed pasteurized milk (0.1% fat content) as shown in Table 1. Sample formulations (Table 1) were selected based on results from previous studies (Bruzzone, Ares, & Giménez, 2013) and preliminary tests in order to obtain yogurts with perceivable differences in their sensory characteristics.

Yogurts were prepared using a Thermomix TM 31 (Vorwerk Mexico S. de R.L. de C.V., Mexico D.F., Mexico). The solid ingredients were mixed with the milk, previously heated to 50 °C. The dispersion was mixed for 1 min under gentle agitation (100 rpm), heated to 90 °C for 5 min and cooled to 42 °C. Then, the mix was placed in 1000 mL glass containers and inoculated with 1 mL of lactics cultures, prepared by dispersing lyophilized cultures of Streptococcus thermophilus, Lactobacillus bulgaricus, Lactobacillus acidophilus, and Bifidobacterium lactis (Yo-Mix 205 LYO 250 DCU, Danisco, France) in UHT skim milk to a concentration of 250 DCU/L. After the addition of cultures the mix was manually agitated for 30 s.

Fermentation was carried out in a temperature controlled oven at 42 ± 1 °C and stopped when the sample reached a pH of 4.55 (after 6 h, depending on the formulation). When the final pH was reached, the coagulum was broken by agitation each yogur for 3 min using a Thermomix TM 31 at 100 rpm. After that, yogurts were placed in 500 mL glass containers, cooled under agitation to 25 °C in a water bath at 5 °C, and then stored refrigerated (5 °C) for 24 h, prior to evaluation.

Samples for all the sensory evaluations performed by trained assessors and consumers were served in plastic containers at 10 °C and coded with 3-digit random numbers and presented following a William’s Latin square design. Twenty grams of yogurt were served for all the evaluations performed by consumers, except for the projective mapping tasks, when 30 g were served to each assessor. For the evaluations performed by the trained assessors, 30 g of yogurt were served.

#### 2.2. Descriptive analysis

The sensory panel consisted of nine assessors, ages ranging from 23 to 48 years old, 66% female. Assessors were selected and trained according to the guidelines of the ISO 8586:2012 standard (ISO, 2012).

In a first session, assessors were presented with four yogurt samples, representing a wide range of sensory characteristics (two commercial samples of plain stirred yogurt, and two formulated yogurts, one with each type of prebiotic component). Assessors were asked to try the yogurts and to individually generate attributes to describe them. Then, through open discussion with the panel leader, assessors agreed on the best attributes to fully describe samples, their definitions and how to evaluate them. The final list of attributes was the following: syneresis, ropiness, thickness, creaminess, roughness, lumpiness, melting, sweetness, sourness, vanilla flavor, milky flavor, sweet aftertaste and sour aftertaste. Definitions and references are shown in Table 2.

In successive sessions, assessors were trained in the quantification of the selected descriptors using unstructured scales. Commercial and formulated yogurts with different sensory characteristics were used during training. A total of fifteen sessions lasting 20 min each were used to train the panel. The sessions were carried out on separate days. Assessors’ performance was checked using PanelCheck® (Tomic et al., 2010).

After the training phase samples were evaluated using 10-cm unstructured line scales anchored with the terms ‘low’ at the left and ‘high’ at the right. Two replications of each sample were evaluated by each assessor. Assessors evaluated four yogurts in each session.

### Table 1

<table>
<thead>
<tr>
<th>Samples</th>
<th>Prebiotic component (%)</th>
<th>Commercial sugar (%)</th>
<th>Stabilizer (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Inulin</td>
<td>4.0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Inulin</td>
<td>4.0</td>
<td>0.075</td>
</tr>
<tr>
<td>3</td>
<td>Inulin</td>
<td>8.0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Inulin</td>
<td>8.0</td>
<td>0.075</td>
</tr>
<tr>
<td>5</td>
<td>FOS</td>
<td>4.0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>FOS</td>
<td>4.0</td>
<td>0.075</td>
</tr>
<tr>
<td>7</td>
<td>FOS</td>
<td>8.0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>FOS</td>
<td>8.0</td>
<td>0.075</td>
</tr>
</tbody>
</table>

* Fructooligosaccharide.
Testing took place in a sensory laboratory in standard sensory booths that were designed in accordance with ISO 8589 (ISO, 2007), under artificial daylight and temperature control (22 °C). Still mineral water was used for rinsing between samples.

2.3. Consumer-based sensory characterization

2.3.1. Consumers and data collection

A total of 243 consumers were recruited from the consumer database of the Food Science and Technology Department of Universidad de la República (Montevideo, Uruguay) based on their consumption of yogurts (at least twice a month), availability and interest to participate. Participants were aged between 18 and 55 years and the percentage of female participants was 61%. The sample comprised varying household compositions, income levels and education levels, but was not representative of the population of Montevideo. Participants gave written informed consent and received a small gift for their participation.

Testing took place in a sensory laboratory in individual sensory booths, designed in accordance with ISO 8589 (2007). Artificial daylight, constant temperature (22 °C) and air circulation were controlled. Still mineral water was available for rinsing.

Participants were randomly divided into three groups of 81 consumers. Each group evaluated the eight yogurt samples using different methodologies: check-all-that-apply questions, projective mapping and polarized sensory positioning. Chi-square tests indicated that differences in the gender, age and yogurt consumption frequency distributions of the groups were not significant.

2.3.2. Check-all-that-apply question

Consumers were asked to complete a check-all-that-apply (CATA) question with 19 terms related to the sensory characteristics of the yogurts: sweet, smooth, fluid, sticky, off-flavor, lumpy, heterogeneous, thick, sour, creamy, milky flavor, rough, liquid, soft, homogeneous, viscous,ropy, consistent, and vanilla flavor. Consumers were asked to check all the terms that they considered appropriate to describe each yogurt. The terms were selected based on published data (Bayarri, Carbonell, Barrios, & Costell, 2011; Bruzzone et al., 2013), considering the descriptors selected by the trained assessors and preliminary studies. Based on recommendations by Ares et al. (2014), the order in which the sensory terms were listed was balanced within and across consumers, following William's Latin Square experimental design.

2.3.3. Projective mapping

Consumers were asked to try the eight samples and to locate them on an A3 white sheet (42 × 30 cm), according to their similarities or dissimilarities. Consumers were explained that they had to complete the task according to their own criteria and that there were no right or wrong answers. They were also explained that two samples close together on the sheet would correspond to very similar samples and that if they perceived two samples as very different they had to locate them very distant from each other. After positioning the samples on the evaluation sheet consumers were asked to provide a brief description of them.

2.3.4. Polarized sensory positioning (PSP)

Three of the eight samples were selected as poles (samples 1, 4 and 7) based on their formulation (Table 1). Each consumer received 30 g of each one of the three poles and approximately 20 g of the eight samples, coded with three-digit random numbers. Consumers were asked to quantify the overall difference between the coded samples and the three different poles using a 10 cm line scale ranging from ‘exactly the same’ to ‘completely different’. For each consumer the distance from ‘exactly the same’ to the mark on the scale was measured for each sample and each of the poles.

2.4. Data analysis

2.4.1. Descriptive analysis

Data from descriptive analysis was analyzed using ANOVA on the scores of each of the 13 attributes, considering sample, session, assessor and their interaction as sources of variation. Performance of the trained panel was considered adequate since interactions of assessor × session, assessor × sample and sample × session were not significant. A 5% significance level was considered in the analyses. When the effects were significant, honestly significant differences were calculated using Tukey’s test.

Principal component analysis (PCA) was performed on the correlation matrix of the attribute scores averaged across assessors for the characteristics that significantly discriminated among samples. Confidence ellipses were constructed using bootstrap techniques (Husson, Le Dîen, & Pagès, 2005).

2.4.2. Check-all-that-apply (CATA) questions

Frequency of use of each one of the terms of the CATA question was determined by counting the number of consumers that used that term
to describe each sample. Cochran’s Q test (Manoukian, 1986) was carried out to identify significant differences among samples for each of the sensory terms.

Correspondence analysis (CA) considering chi-square distance was carried out on the matrix containing the frequency of use of each term for each sample. Confidence ellipses around the projected coordinates of the samples were obtained using bootstrapping (Ringrose, 2012).

2.4.6. Comparison of the methodologies

Were computed over subsets for each number of consumers. (Robert & Escouffier, 1976). It takes the value of 0 if the configurations are uncorrelated and the value of 1 if the configurations are homothetic. A permutation test was used to evaluate the significance of the RV coefficient (Josse, Pagès, & Husson, 2008).

All statistical analyses were performed using R language (R Core Team, 2013). FactoMineR (Lê, Josse, & Husson, 2008) was used to perform MFA and PCA and to calculate RV coefficients. SensoMineR was used to obtain confidence ellipses from descriptive analysis (Lê & Husson, 2008), and cabootcrs used to perform CA and obtain confidence ellipses (Ringrose, 2012).

3. Results

3.1. Descriptive analysis data

Significant differences (p < 0.05) among samples were identified for all sensory attributes. As shown in Table 3, differences among samples in the intensity of the attributes ranged between 2.4 and 4.8 points in the 10-cm unstructured scale. The smallest differences were identified in the attributes syneresis, melting and milky flavor.

The first and second dimensions of the PCA of descriptive analysis data explained 89.7% of the variance. As shown in Fig. 1(a), the first dimension sorted samples according to their sugar concentration. Samples 1, 2, 5, and 6 (formulated with 4.0% sugar) were located to the right of the first dimension and were characterized by their sourness, sour aftertaste and milky flavor. On the contrary, samples formulated with 8.0% sugar (samples 3, 4, 7 and 8) were located at the right of the first dimension and were mainly characterized by their sweetness, sweet aftertaste and ripiness. Samples were sorted in the second dimension of the PCA according to their stabilizer concentration. Samples containing 0.075% of stabilizer (2, 4, 6, and 8) were located at positive values of the second dimension, being characterized by their high roughness, lumpiness and syneresis scores. Meanwhile, samples formulated without a stabilizer (1, 3, 5 and 7) were located at negative values of the second dimension and were characterized by their thickness and creaminess (Fig. 1(a)). Samples with different prebiotic components were not significantly discriminated in the first two dimensions of the PCA since their confidence ellipses overlapped (e.g. samples 2 and 6 and samples 4 and 8).

The third dimension of the PCA, which explained 7.77% of the variance, sorted samples according to their prebiotic ingredient. Samples formulated with inulin tended to be located at positive values of the third dimension, while samples formulated with fructooligosaccharide were located at negative values of the third dimension (Fig. 1(b)). Samples with inulin mainly differed with those formulated with FOS in their thickness and roughness.

3.2. Consumer-based sensory characterizations

3.2.1. Check-all-that-apply questions

Significant differences among samples were found in the frequency with which consumers used 15 of the 19 terms included in the CATA

Table 3

Average scores for the sensory attributes of the eight yogurt samples evaluated using descriptive analysis.

<table>
<thead>
<tr>
<th>Samples</th>
<th>Syneresis</th>
<th>Ropiness</th>
<th>Consistency</th>
<th>Creaminess</th>
<th>Roughness</th>
<th>Lumpiness</th>
<th>Melting</th>
<th>Sweet</th>
<th>Sour</th>
<th>Vanilla flavor</th>
<th>Milky flavor</th>
<th>Sweet aftertaste</th>
<th>Sour aftertaste</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6b</td>
<td>2.1bc</td>
<td>5.5b</td>
<td>6.6b</td>
<td>1.2ab</td>
<td>0.7ab</td>
<td>5.5bc</td>
<td>2.5b</td>
<td>6.0a</td>
<td>1.6b</td>
<td>3.9ab</td>
<td>0.8b</td>
<td>5.1a</td>
</tr>
<tr>
<td>2</td>
<td>2.5b</td>
<td>1.0ef</td>
<td>5.3b</td>
<td>3.6b</td>
<td>5.1a</td>
<td>3.6b</td>
<td>6.8b</td>
<td>2.2b</td>
<td>6.0a</td>
<td>1.3b</td>
<td>4.0b</td>
<td>1.3b</td>
<td>4.5a</td>
</tr>
<tr>
<td>3</td>
<td>3.0c</td>
<td>4.8a</td>
<td>7.4a</td>
<td>7.0a</td>
<td>1.7a</td>
<td>0.5ab</td>
<td>4.9b</td>
<td>6.0b</td>
<td>2.4b</td>
<td>5.2b</td>
<td>1.5b</td>
<td>4.7b</td>
<td>1.3b</td>
</tr>
<tr>
<td>4</td>
<td>1.9abc</td>
<td>2.4bcd</td>
<td>5.1b</td>
<td>5.6b</td>
<td>2.3b</td>
<td>1.9bc</td>
<td>5.4bc</td>
<td>6.6b</td>
<td>2.1b</td>
<td>4.9b</td>
<td>2.5bcd</td>
<td>4.9b</td>
<td>1.7b</td>
</tr>
<tr>
<td>5</td>
<td>0.6a</td>
<td>1.4def</td>
<td>4.7b</td>
<td>5.1bcd</td>
<td>1.0bc</td>
<td>0.2a</td>
<td>7.4a</td>
<td>3.0b</td>
<td>5.4b</td>
<td>1.7b</td>
<td>1.5bcd</td>
<td>0.9b</td>
<td>4.4a</td>
</tr>
<tr>
<td>6</td>
<td>3.0c</td>
<td>0.7c</td>
<td>2.7c</td>
<td>2.8c</td>
<td>2.6b</td>
<td>3.2ab</td>
<td>7.4b</td>
<td>3.1b</td>
<td>5.8b</td>
<td>1.8b</td>
<td>4.2b</td>
<td>0.7b</td>
<td>5.0b</td>
</tr>
<tr>
<td>7</td>
<td>0.6a</td>
<td>3.9ab</td>
<td>5.7b</td>
<td>6.5b</td>
<td>0.3c</td>
<td>0.1a</td>
<td>4.7b</td>
<td>6.9b</td>
<td>2.3b</td>
<td>5.3b</td>
<td>2.3cd</td>
<td>5.3</td>
<td>1.5a</td>
</tr>
<tr>
<td>8</td>
<td>1.9ab</td>
<td>3.6bc</td>
<td>4.9b</td>
<td>4.4abc</td>
<td>1.9bc</td>
<td>0.9cd</td>
<td>5.8b</td>
<td>6.3b</td>
<td>1.9b</td>
<td>4.7b</td>
<td>1.8ab</td>
<td>5.1a</td>
<td>1.4a</td>
</tr>
</tbody>
</table>

Attributes were evaluated in a 10 cm unstructured intensity scale; different superscript letters within a row imply a significant difference according to Tukey's test (p < 0.05).
question (Table 4). This result suggests that the CATA question was able to detect differences in consumers’ perception of the sensory characteristics of the evaluated yogurts. The only terms in which differences among samples were not significant were fluid, milky flavor, viscous and ropy.

The first two dimensions of the correspondence analysis explained 88.5% of the inertia of the experimental data. As shown in Fig. 2 samples were sorted along the first dimension according to their sugar content. Samples formulated with 8.0% sucrose (3, 4, 7 and 8) were located at negative values of the first dimension, being described with the terms sweet, vanilla flavor and smooth. Samples formulated with 4.0% sucrose were located at positive values of the first dimension and were mainly characterized by their off-flavor and sourness. These samples were sorted into two groups. Samples 1 and 5, formulated without a stabilizer, were described as thick and consistent (Fig. 2), whereas samples containing 0.075% of stabilizer (samples 2 and 6) were described as rough, heterogeneous, and liquid.

The third and fourth dimensions of the CA explained 9.0% of the inertia of the experimental data but did not provide relevant information about the sensory characteristics of the samples (data not shown).

3.2.2. Projective mapping

The first two dimensions of the MFA explained 44.6% of the variance, while the first four dimensions explained 71.4%. As shown in Fig. 3(a), the first dimension sorted samples according to their sucrose concentration. Samples formulated with 8.0% sucrose (yogurts 3, 4, 7, and 8) were described as sweet, very sweet, thick, andasty, whereas samples formulated with 4.0% (formulations 1, 2, 5, and 6) were described as sour, very sour, not very sweet, bitter, natural and disgusting. This last group of four samples were located in the second dimension of the MFA according to their stabilizer concentration. Samples formulated without a stabilizer (samples 1 and 5) were located at positive values of the second dimension. Meanwhile, samples containing 0.075% of stabilizer were located at negative values of this dimension, being described as not very thick, rough and heterogeneous.

The third and fourth dimensions of the MFA sorted samples according to their prebiotic component. Samples formulated with inulin (1, 2, 3 and 4) were located at positive values of the third dimension. These samples were mainly described as disgusting (Fig. 3(b)). Meanwhile, samples formulated with fructooligosaccharide (samples 5, 6, 7 and 8) were located at negative values of the third dimension. The fourth dimension sorted samples according to their stabilizer concentration (Fig. 3). Samples 5 and 7, formulated without a stabilizer, were mainly described with the terms vanilla flavor, natural and thick.

3.2.3. Polarized sensory positioning

The first and second dimensions of the MFA explained 60.6% of the variance of the experimental data. Samples were sorted into two main groups according to their sucrose concentration. Samples 1, 2, 5 and 6, formulated with 4.0% sucrose were located at positive values of the
Table 4

<table>
<thead>
<tr>
<th>Term</th>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweet***</td>
<td>5</td>
<td>9</td>
<td>70</td>
<td>65</td>
<td>16</td>
<td>14</td>
<td>73</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Smooth*</td>
<td>21</td>
<td>12</td>
<td>26</td>
<td>31</td>
<td>22</td>
<td>17</td>
<td>31</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Fluid†</td>
<td>21</td>
<td>19</td>
<td>20</td>
<td>28</td>
<td>14</td>
<td>32</td>
<td>19</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Sticky*</td>
<td>11</td>
<td>19</td>
<td>12</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Off-flavor***</td>
<td>28</td>
<td>42</td>
<td>14</td>
<td>9</td>
<td>20</td>
<td>23</td>
<td>7</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Lumpy***</td>
<td>5</td>
<td>25</td>
<td>2</td>
<td>10</td>
<td>4</td>
<td>15</td>
<td>0</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Heterogeneous***</td>
<td>5</td>
<td>23</td>
<td>2</td>
<td>12</td>
<td>11</td>
<td>19</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Thick***</td>
<td>36</td>
<td>27</td>
<td>19</td>
<td>15</td>
<td>35</td>
<td>19</td>
<td>22</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Sour***</td>
<td>56</td>
<td>59</td>
<td>19</td>
<td>20</td>
<td>48</td>
<td>48</td>
<td>9</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Creamy*</td>
<td>52</td>
<td>27</td>
<td>47</td>
<td>46</td>
<td>49</td>
<td>19</td>
<td>62</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Milky flavor***</td>
<td>22</td>
<td>22</td>
<td>19</td>
<td>28</td>
<td>38</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Rough***</td>
<td>12</td>
<td>27</td>
<td>10</td>
<td>14</td>
<td>15</td>
<td>28</td>
<td>5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Liquid***</td>
<td>9</td>
<td>19</td>
<td>21</td>
<td>20</td>
<td>9</td>
<td>35</td>
<td>17</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Soft*</td>
<td>32</td>
<td>21</td>
<td>49</td>
<td>43</td>
<td>42</td>
<td>33</td>
<td>52</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Homogeneous***</td>
<td>38</td>
<td>26</td>
<td>51</td>
<td>38</td>
<td>42</td>
<td>38</td>
<td>53</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Viscous***</td>
<td>15</td>
<td>21</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>12</td>
<td>9</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Ropy***</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>9</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Consistent***</td>
<td>26</td>
<td>16</td>
<td>20</td>
<td>15</td>
<td>31</td>
<td>11</td>
<td>25</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Vanilla flavor***</td>
<td>7</td>
<td>6</td>
<td>33</td>
<td>28</td>
<td>14</td>
<td>7</td>
<td>40</td>
<td>27</td>
<td></td>
</tr>
</tbody>
</table>

*** Indicates significant differences at \( p < 0.001 \).
** Indicates significant differences at \( p < 0.01 \).
* Indicates significant differences at \( p < 0.05 \).
††† Indicates no significant differences \( p > 0.05 \) according to Cochran's Q test.

first dimension, as shown in Fig. 4. On the other hand, samples formulated with 8.0% sucrose were sorted into two groups. Yogurts 4 and 8, formulated with 0.075% of a stabilizer were located at positive values of the second dimension, whereas samples formulated without a stabilizer were located at negative values of the second dimension.

The third and fourth dimensions of the MFA explained 18.83% of the variance but did not provide useful information for explaining the influence of formulation on the sensory characteristics of the samples. The position of samples on these two dimensions (data not shown) was not explained by differences in their formulation.

3.2.4. Stability of sample configurations

The stability of the sample configurations was determined using a bootstrapping resampling approach. As shown in Fig. 5, the average RV coefficient of sample configurations rapidly increased as the number of consumers in the virtual panel increased, while standard deviations decreased. Sample configurations from projective mapping were the least stable, regardless of the number of consumers in the virtual panel. PSP provided more stable sample configurations than CATA questions when the number of consumers in the virtual panel was lower than 35. However, when the number of consumers was higher than 35 the stability of the sample configurations from both methodologies did not differ (Fig. 5).

Blancher et al. (2012) considered an average RV coefficient of 0.95 as an indicator of stable sample configurations. Considering this threshold, sample configurations from the PSP and CATA were stable. The number of consumers necessary to reach stable sample configurations was 25 for PSP and 35 for CATA questions. However, sample configurations from projective mapping did not reach stability since the maximum average coefficient corresponded to 0.94 (lower than 0.95). If an average RV coefficient of 0.90 is considered as a stability criterion (Vidal et al., 2014), sample configurations from projective mapping can also be regarded as stable and the minimum number of consumers necessary to reach stability corresponded to 80.

3.3. Comparison of the methodologies

3.3.1. Comparison of sample configurations

The RV coefficient was used to compare sample configurations from the four methodologies used for sensory characterization of yogurt samples. As shown in Table 5, sample configurations obtained using CATA questions showed the highest similarity with descriptive analysis (DA) with trained assessors (0.93), followed by PSP (0.86). On the contrary, sample configurations from projective mapping and DA showed the lowest RV coefficient (0.54). Although both projective mapping and PSP are based on the evaluation of global differences among samples, the RV between sample configurations from these methodologies was low (0.61) (Table 5).

3.3.2. Sample descriptions

Three of the four methodologies provided a description of the sensory characteristics responsible for similarities and differences among samples. The terms elicited in the description phase of projective mapping were similar to the attributes selected by the trained assessors for descriptive analysis and to the terms included in the CATA question. According to the three methodologies, the sensory characteristics responsible for the largest differences among samples were related to sweetness and sourness. These were the most frequently mentioned categories in the description phase of the projective mapping task. Other flavor and texture characteristics were also frequently mentioned, such as creamy, thick, tastelass and smooth. Besides, consumers also elicited hedonic terms, such as delicious and disgusting. It is important to highlight that terms related to aftertaste, ropiness and syneresis were
not elicited by consumers (Fig. 3) although they were used by the trained assessors (Table 2).

4. Discussion

Descriptive analysis (DA) with trained assessors provided a quantitative measure of the intensity of 13 sensory characteristics of the yogurts and enabled the identification of significant differences among samples in each of the evaluated characteristics (Table 3). DA showed that sucrose and stabilizer concentrations were responsible for the main differences in the sensory characteristics of the samples, whereas the type of prebiotic ingredient did not have a large impact on the sensory characteristics of the samples (Fig. 1). In the present work DA provided more accurate and detailed information than rapid methodologies, having a higher discriminating ability to detect significant differences among samples. Similar results have been reported by other authors (Cartier et al., 2006; Dehlholm, Brockhoff, Meinert, Aaslyng, & Bredie, 2012).

Although DA provided accurate and detailed information, it should be taken into account that almost two months were needed to train the panel to get reliable information about the sensory characteristics of the yogurts. In the present work three rapid methodologies for sensory characterization were carried out by naïve consumers, which markedly reduced the time needed for their implementation. Each of these methodologies were implemented in two sessions of 6 h, which consists of a considerable reduction of time.

The three methodologies provided similar information, sorting samples according to their sucrose concentration into two markedly different groups. However, they differed in their ability to detect differences among samples due to other formulation variables.

Although several authors have reported that sorting and projective mapping are less discriminating than methodologies based on the evaluation of specific sensory attributes (Cruz et al., 2013; Dehlholm, Brockhoff, Meinert, et al., 2012; Moussaoui & Varela, 2010; Veinand, Godefroy, Adam, & Delarue, 2011), the trend was not found in the present work. The CATA question and PSP showed a lower discriminative ability than projective mapping. The former two methodologies allowed identifying differences in the sensory characteristics of samples due to stabilizer concentration only among samples formulated with one of the levels of sucrose. CATA questions were not able to detect differences among samples formulated with 8.0% sucrose (Fig. 2), whereas PSP only identified differences among samples with different stabilizer concentrations at the lowest sucrose concentration (Fig. 4). The discriminative capacity of projective mapping was higher than that of CATA and PSP. Samples formulated with 8.0% sucrose were sorted according to their stabilizer concentration. However, samples with 4.0% were sorted in two groups that did not correspond to differences in their formulation. Besides, in the third and fourth dimensions samples were sorted according to the prebiotic ingredient included in the formulation.
Regarding the similarity between DA and rapid methodologies, sample configurations from the CATA question were the most similar (Table 5). This can be explained by the fact that the CATA question is based on the evaluation of a specific product’s sensory characteristics, similarly to DA. The high similarity between results from CATA questions and DA have been reported by other authors for the evaluation of different dairy products (Ares, Barreiro, Deliza, Giménez, & Gámbaro, 2010; Bruzzzone et al., 2012; Dooley et al., 2010; Reinbach, Giacalone, Ribeiro, Bredie, & Frøst, 2014).

Sample configurations from projective mapping were the least similar to those from DA (Table 5). Low RV coefficients between sample configurations from projective mapping and descriptive mapping have been reported by other authors. Dehholm, Brockhoff, Meinert, et al. (2012) reported that the RV coefficient between DA and projective mapping was 0.48 when evaluating liver paté. In this sense it should be highlighted that differences between results from projective mapping and DA do not mean lack of validity. Projective mapping encourages a generation of a synthetic representation of the products, which is usually inherited when assessors are asked to focus their attention on multiple attributes (Prescott, 1999; Small & Prescott, 2005). For this reason, sample configurations depend on the relative importance that assessors give to the sensory characteristics of the samples, providing different information than attribute-based methodologies.

The reliability of sensory characterizations from DA can be analyzed by monitoring the global performance of the panel, as well as the performance of each individual assessor. Considering that assessors are trained in attribute identification and scaling, the dispersion of the scores provided to each attribute for each sample is used to estimate panel agreement. Besides, samples are evaluated in duplicate or triplicate, which enables the analysis of global and individual reproducibility (Lawless & Heymann, 2010). However, rapid methodologies for sensory characterization do not require training and are usually performed in a unique session, which makes it difficult to evaluate their reliability.

Although no standard procedure is available for evaluating the reliability of sensory characterizations obtained with these methodologies, several approaches have been used. Faye et al. (2006) and Blancher et al. (2012) proposed to estimate the reliability of sample configurations using simulated repeated experiments through a bootstrapping resampling approach. Blancher et al. (2012) argued that a sorting map could be considered stable if sampling repeatedly from the population of interest provides equivalent sorting maps. Using this approach CATA questions and PSP proved to be highly reliable, providing sample configurations that reach average RV coefficients higher than 0.95. Projective mapping was less stable than the other two methodologies and did not reach an average RV value of 0.95. The minimum number of consumers needed to reach stable sample configurations using CATA questions and projective mapping are similar to those reported by Ares, Tárrega, Izquierdo, and Jaeger (2014) and Vidal et al. (2014).

Differences in the stability of sample configurations from the three rapid methodologies can be explained considering how assessors are asked to evaluate samples. In PSP assessors have to evaluate similarities and differences between samples and the poles, whereas in CATA questions they used a predetermined list of terms to describe samples. When projective mapping is used, each assessor selects his/her own criteria for describing samples. In this sense it should be highlighted that differences between results from projective mapping and DA do not mean lack of validity. Projective mapping encourages a generation of a synthetic representation of the products, which is usually inherited when assessors are asked to focus their attention on multiple attributes (Prescott, 1999; Small & Prescott, 2005). For this reason, sample configurations depend on the relative importance that assessors give to the sensory characteristics of the samples, providing different information than attribute-based methodologies.

### Table 5
RV coefficients between sample configurations in the first two dimensions of multivariate statistical techniques for the four methodologies used for sensory characterization of yogurt samples.

<table>
<thead>
<tr>
<th></th>
<th>QDA</th>
<th>CATA</th>
<th>Projective mapping</th>
<th>PSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>QDA</td>
<td>1</td>
<td>0.93</td>
<td>0.54</td>
<td>0.86</td>
</tr>
<tr>
<td>CATA</td>
<td>–</td>
<td>1</td>
<td>0.64</td>
<td>0.87</td>
</tr>
<tr>
<td>Projective mapping</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>0.61</td>
</tr>
<tr>
<td>PSP</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
</tbody>
</table>

The reliability of sensory characterizations from DA can be analyzed by monitoring the global performance of the panel, as well as the performance of each individual assessor. Considering that assessors are trained in attribute identification and scaling, the dispersion of the scores provided to each attribute for each sample is used to estimate panel agreement. Besides, samples are evaluated in duplicate or triplicate, which enables the analysis of global and individual reproducibility (Lawless & Heymann, 2010). However, rapid methodologies for sensory characterization do not require training and are usually performed in a unique session, which makes it difficult to evaluate their reliability.
perception during new product development. Also, consumers described samples using sensory and hedonic terms, in agreement with other studies (Ares, Varela, Rado, & Giménez, 2011; Veinand et al., 2011). Although the use of hedonic terms can be regarded as a limitation of consumer-based sensory characterizations, it should be taken into account that hedonic information can be used to identify relevant sensory characteristics for consumers for the design of marketing and communication strategies.

Finally, it is worth mentioning that PSP has the advantage of enabling the comparison of sample configurations from different sessions, which cannot be easily achieved using the other two rapid methodologies. This advantage is interesting considering the iterative nature of new product development.

5. Conclusions

Traditional descriptive analysis is still the most robust methodology for sensory characterization, providing more detailed and accurate information. However, rapid methodologies provide the advantage of providing information about the sensory characteristics of products in short time frames, which can be a major advantage in the first steps of product development. Also, sensory characterization is performed from a consumer-based point of view, which may contribute to the development of more successful products. In the present work check-all that-apply (CATA) questions, projective mapping and polarized sensory positioning provided similar information than descriptive analysis for sensory characterization of functional yogurts. However, DA provided positioning provided similar information than descriptive analysis for type of product, CATA questions provided the most similar information for sensory characterization, providing more detailed and accurate information about the sensory characteristics of products in other studies (Ares, Varela, Rado, & Giménez, 2011; Veinand et al., 2011).

Acknowledgements

The authors are indebted to CAPES-Brasil and to Comisión Sectorial de Investigación Científica (Universidad de la República, Uruguay) for financial support.

References


