Preference mapping of dulce de leche commercialized in Brazilian markets

L. V. Gaze,*¹ B. R. Oliveira,*¹ L. L. Ferrão,† D. Granato,‡ R. N. Cavalcanti,§ C. A. Conte Júnior,*¹ A. G. Cruz,# and M. Q. Freitas*¹
*¹Universidade Federal Fluminense (UFF), Faculdade de Veterinária, Pós-Graduação em Higiene Veterinária e Processamento de Produtos de Origem Animal, 24230-340 Niterói, Rio de Janeiro, Brazil
†Universidade Federal Rural do Rio de Janeiro (UFRJR), Instituto de Tecnologia, Pós-Graduação em Ciência e Tecnologia de Alimentos, BR 465 km 7, 23890-000 Seropédica, Rio de Janeiro, Brazil
‡Universidade Estadual de Ponta Grossa (UEPG), Departamento de Engenharia de Alimentos, Av. Carlos Cavalcanti, 4748, 84030-900, Ponta Grossa, Brazil
§Universidade Estadual de Campinas (UNICAMP), Faculdade de Engenharia de Alimentos, Av. Carlos Cavalcanti, 4748, 84030-900, Ponta Grossa, Brazil
#Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro (IFRJ), Equipe de Alimentos, Rua Senador Furtado, 171/175, Maracanã, 20270-021 Rio de Janeiro, Brazil

ABSTRACT

Dulce de leche samples available in the Brazilian market were submitted to sensory profiling by quantitative descriptive analysis and acceptance test, as well sensory evaluation using the just-about-right scale and purchase intent. External preference mapping and the ideal sensory characteristics of dulce de leche were determined. The results were also evaluated by principal component analysis, hierarchical cluster analysis, partial least squares regression, artificial neural networks, and logistic regression. Overall, significant product acceptance was related to intermediate scores of the sensory attributes in the descriptive test, and this trend was observed even after consumer segmentation. The results obtained by sensometric techniques showed that optimizing an ideal dulce de leche from the sensory standpoint is a multidimensional process, with necessary adjustments on the appearance, aroma, taste, and texture attributes of the product for better consumer acceptance and purchase. The optimum dulce de leche was characterized by high scores for the attributes sweet taste, caramel taste, brightness, color, and caramel aroma in accordance with the preference mapping findings. In industrial terms, this means changing the parameters used in the thermal treatment and quantitative changes in the ingredients used in formulations. Key words: sensometric technique, dulce de leche, optimization

INTRODUCTION

Multivariate statistical techniques are very useful for analyzing complex data obtained from consumer and trained panelists, as they are able to provide a spatial representation of the experimental data (Ennis and Ennis, 2013; Zielinski et al., 2014). In addition, they proportionate the acquisition of the drivers of consumer liking and purchase intent, which is a concept frequently used in sensory and consumer research (Bi, 2012). Indeed, when differences are noted among food products that may drive consumer liking and purchase intent, it is helpful to visualize a space in which each possible product is placed (Cadena et al., 2012). Another outcome is the consumer’s food choices, which are complex interactions between the product and the consumer, including the consumer’s personality and previous familiarity and experience with the product (Cadena et al., 2013).

Dulce de leche (DL) is a dairy-based product produced from the dehydration of fluid milk under conditions of temperature and pressure, which vary according to the manufacturer, with added sucrose. Sodium bicarbonate can also be added to the milk to increase the typical brown color caused by the Maillard reaction, which is favored at pH 6 to 7, and to avoid the formation of lumps that result from the concentration of acidic compounds generated during evaporation (Perrone et al., 2011). Dulce de leche is manufactured and consumed in Latin America, especially Argentina and Uruguay and to a lesser extent in Chile, Paraguay, Brazil, and Bolivia (Zalazar and Perotti, 2011). In Brazil, production of DL is performed exclusively by small dairy producers, without control of the operational parameters, resulting in products with different physicochemical properties, which have a direct influence on the sensory characteristics. The consumer market is essentially formed from 2 groups: end consumers and institutional consumers (food service and food industry). The institutional segment is a large market, represented by fast-food chains, prepared (ready-to-eat) foods, ice cream shops, confectioners, and wholesale centers. In
addition, hotels, motels, and convenience stores have also acquired DL in the form of mini portions.

A food product’s quality depends on continuous consumer satisfaction with respect to the sensory attributes, thus it is necessary to evaluate the relevant attributes for acceptance and purchase intent of a food product. In this context, our study aimed to identify potential for improvement of DL available in the Brazilian market and the relevant sensory attributes for acceptance and purchase intent. Sensometric techniques were used to analyze the results, such as partial least squares regression, artificial neural networks, and logistic regression. In addition, external preference mapping was performed using sensory profiling and the acceptance values obtained in the consumer test.

The results of this study may provide useful information for dairy companies in the development and marketing of new products or the reformulation of existing products, in addition to establishing quality control specifications.

**MATERIALS AND METHODS**

**Samples**

Seven commercial brands of dulce de leche were purchased from dairy companies inspected by the Brazilian Federal Inspection Service (SIF), which enables marketing throughout the Brazilian territory. The samples were acquired immediately after processing so that shelf life was not considered as a factor in this study. Samples were coded as I, II, III, IV, V, VI, and VII to maintain the confidentiality of the trademarks and manufacturers. The ingredients used in manufacturing each product, as well as the nutritional composition displayed on the label, are shown in Table 1. The samples were presented in randomized and balanced manner for both the descriptive analysis and consumer tests, according to MacFie et al. (1989).

Although the first impression suggests sampling only a limited number of products, it is important to note that the Brazilian DL industry is still largely handmade, without continuous processing, and it is time consuming and energy intensive (Perrone et al., 2014). All these factors contribute to products in the local market having considerable variation with respect to color, consistency, body, texture, composition, and presentation. In fact, physicochemical and instrumental quality parameters were recently evaluated using the same sample size (Gaze et al., 2015).

**Sensory Profiling**

To quantify the descriptive attributes of the samples, quantitative descriptive analysis (QDA) was

---

**Table 1. Ingredients and nutritional composition of dulce de leche samples**

<table>
<thead>
<tr>
<th>Brand</th>
<th>Main ingredients</th>
<th>Saturated fat (g/100 g)</th>
<th>Trans fat (g/100 g)</th>
<th>Sodium (mg/100 g)</th>
<th>Total fat (g/100 g)</th>
<th>Calcium (g/100 g)</th>
<th>Protein (g/100 g)</th>
<th>Total carbo (g/100 g)</th>
<th>Energy (kcal/100 g)</th>
<th>Fat (% of energy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Milk, sugar, sodium bicarbonate</td>
<td>0.280</td>
<td>—</td>
<td>0.115</td>
<td>0.175</td>
<td>0.175</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>Milk, sugar, sodium bicarbonate, sodium chloride, lactose</td>
<td>0.165</td>
<td>0.025</td>
<td>0.482</td>
<td>0.955</td>
<td>0.955</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>Milk, sugar, sodium bicarbonate, potassium sorbate</td>
<td>0.275</td>
<td>—</td>
<td>0.165</td>
<td>0.325</td>
<td>0.325</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>Milk, sugar, modified starch, sodium citrate, lactase, potassium sorbate</td>
<td>0.165</td>
<td>0.025</td>
<td>0.325</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>Milk, sugar, sodium bicarbonate, potassium sorbate</td>
<td>0.115</td>
<td>—</td>
<td>0.165</td>
<td>0.325</td>
<td>0.325</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>Milk, sugar, sodium bicarbonate, lactose</td>
<td>0.280</td>
<td>—</td>
<td>0.115</td>
<td>0.175</td>
<td>0.175</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VII</td>
<td>Milk, sugar, modified starch, sodium bicarbonate, potassium sorbate</td>
<td>0.165</td>
<td>0.025</td>
<td>0.482</td>
<td>0.955</td>
<td>0.955</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No information provided on the label.
Consumer Test

The consumer tests were carried out under laboratory conditions, at the Sensory Analysis Laboratory of the Universidade Federal Fluminense (UFF, Rio de Janeiro, Brazil), with samples served in plastic cups in a randomized and balanced manner. Crackers and mineral water were provided for cleansing the palate between the samples. One hundred twenty-five consumers (37 men and 88 women) participated in the trial; they were aged between 15 and 57 yr and were regular consumers of dulce de leche (at least once a week).

Before the sensory tests, assessors completed and signed a written consent to participate in the tests, stating that they were physically fit to do so. Acceptance testing was conducted using a 9-point hedonic scale (9 = like extremely, 1 = dislike extremely) to assess the following attributes: appearance, aroma, taste, texture, and overall acceptance; intent to purchase was assessed by a 5-point scale (1 = definitely would not buy, 5 = definitely would buy; Cruz et al., 2012). The attributes color, sweet taste, caramel taste, and consistency were evaluated using a just-about-right (JAR) scale of 9 points, with 9 = extremely over the ideal; 5 = ideal, and 1 = extremely less than the ideal (Chollet et al., 2013; Morais et al., 2014b).

Overview of Sensometric Techniques

Partial least squares regression (PLS) is a method widely used in sensory analysis using both consumer and sensory descriptive analysis by a trained panel, enabling the identification of positive and negative attributes that contribute to product acceptance (Cadena et al., 2012). The PLS components are created by the explanatory variable (X-data) and dependent variable (Y-data), being the original data decomposed into principal components called factors, indicating better performance when the data show a linear behavior (Cruz et al., 2011).

Artificial neural network (ANN) is a mathematical algorithm capable of relating input and output parameters through training, without prior knowledge on the relationships between the process variables (Cevoli et al., 2011). Artificial neural networks require a training process in which the weights of those connections are adjusted, building a model that will allow prediction of the desired parameters (either qualitative or quantitative; Cetó et al., 2013).

Penalty (mean-drop) analysis is used by market researchers and product developers to understand the product attributes that most affect liking, measuring the changes in product liking (or any other measure) due to the product having “too much” or “too little” of a specific attribute (Market Research World, 2013). A few studies have used penalty analysis in a dairy foods consumer study: in cottage cheese (Drake et al., 2011) and in vanilla yogurt with stevia (Narayanan et al., 2013).

Logistic regression (LR) is a powerful tool for the statistical analysis of categorical data with the main purpose to discover exploratory variables that drive a choice outcome, without necessity of evaluating integrals (Ennis and Ennis, 2013). It can be used to cor-
Table 2. Sensory profiling of the dulce de leche samples

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Appearance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color</td>
<td>Print feature of sweet milk generated from the reflectance of light on the product and generating color ranging from yellow to brown.</td>
<td>Beige = dulce de leche (Ylagam) Brown = dulce de leche (Macuco)</td>
</tr>
<tr>
<td>Brightness</td>
<td>Intensity of light reflection.</td>
<td>Low = dulce de leche (Macuco) High = dulce de leche (Além)</td>
</tr>
<tr>
<td>Adhesiveness</td>
<td>Ability of the product to adhere to the spoon or spatula when lying on the surface of the product, using only its own weight, and raised manually up to 90°.</td>
<td>Low = dulce de leche (Além) High = dulce de leche (Macuco; SIF 614)</td>
</tr>
<tr>
<td>Viscosity</td>
<td>Force required to mix the sample with circular motion with a spoon or spatula.</td>
<td>Low = dulce de leche (Além; SIF 202) High = dulce de leche (Macuco; SIF 614)</td>
</tr>
<tr>
<td><strong>Aroma</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heated milk</td>
<td>Flavor of cow milk associated with milk heated to scalding point.</td>
<td>None = pasteurized milk (Sabor da Serra) Strong = 300 mL of pasteurized milk (Sabor da Serra; SIE/RJ 682) heated in 1 L of water in a warm bath for 10 min</td>
</tr>
<tr>
<td>Caramel</td>
<td>Aroma of heated carbohydrates in solutions of concentrated milk.</td>
<td>None = pasteurized milk (Sabor da Serra) Strong = one portion of Toffee Arcólor essence diluted in distilled water at a concentration of 0.12% and the other part of pasteurized milk (Sabor da Serra; SIE/RJ 682)</td>
</tr>
<tr>
<td><strong>Taste</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweet</td>
<td>Perceived taste on the tongue stimulated by sugars.</td>
<td>Weak = condensed milk (Moça) and pasteurized milk (Sabor da Serra) at a proportion of 1:3 (wt/wt) Strong = condensed milk (Moça) and pasteurized milk (Sabor da Serra) at a proportion of 3:1 (wt/wt)</td>
</tr>
<tr>
<td>Heated milk</td>
<td>Flavor of cow milk heated up to its scalding point.</td>
<td>None = pasteurized milk (Sabor da Serra) Strong = 300 mL of pasteurized milk (Sabor da Serra) heated in 1 L of water in a warm bath for 10 min</td>
</tr>
<tr>
<td>Caramel</td>
<td>Taste of carbohydrates heated in solutions of concentrated milk.</td>
<td>None = pasteurized milk (Sabor da Serra) Strong = one portion of Toffee Arcólor essence diluted in distilled water at a concentration of 0.12% and the other part of pasteurized milk (Sabor da Serra)</td>
</tr>
<tr>
<td><strong>Butter</strong></td>
<td>Flavor associated with the unsalted butter.</td>
<td>None = creamy dulce de leche (Macuco) Strong = glucose corn syrup (Karô) and pasteurized milk (Sabor da Serra) at a proportion of 3:5 (wt/wt)</td>
</tr>
<tr>
<td>Pungent aftertaste</td>
<td>Slightly spicy sensation after swallowing</td>
<td>None = glucose corn syrup (Karô) and pasteurized milk (Sabor da Serra) at a proportion of 3:5 (wt/wt) High = glucose corn syrup (Karô) (Unilever)</td>
</tr>
<tr>
<td>Oil layer</td>
<td>Feeling oily film on the oral mucosa and lips.</td>
<td>Low = dulce de leche - Souvenir High = unsalted butter (Aviação) and dulce de leche (Ylagam) at a proportion of 1:5 (wt/wt)</td>
</tr>
<tr>
<td><strong>Texture</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adhesiveness</td>
<td>Adhesion of the product on the palate when pressed against the roof of the mouth by the tongue.</td>
<td>Low = dulce de leche (Além) High = dulce de leche (Macuco; SIF 614)</td>
</tr>
<tr>
<td>Viscosity</td>
<td>Speed in which the sample crumbles in the mouth; the greater the speed, the less viscous the product is.</td>
<td>Low = dulce de leche (Além) High = dulce de leche (Macuco)</td>
</tr>
<tr>
<td>Sandiness</td>
<td>Presence of small granules dispersed uniformly throughout the sample.</td>
<td>None = dulce de leche (Além) High = dulce de leche (Itambé)</td>
</tr>
</tbody>
</table>

1Ylagam, São Paulo, Brazil; Macuco, Rio de Janeiro, Brazil; Além, São Paulo, Brazil; Sabor da Serra, São Paulo, Brazil; Moca, São Paulo, Brazil; Itaocara, São Paulo, Brazil; Aviação, São Paulo, Brazil; Unilever, São Paulo, Brazil; Itambé, Minas Gerais, Brazil.
relate affective tests, including purchase intent (which, in this case, is presented as a dichotomous response variable), providing information about the most important attributes at the time of purchase (Prinyawiwatkul and Chompreeda, 2007).

**Preference Mapping**

External preference mapping consists of a regression where consumer acceptability (hedonic) ratings are the response (dependent) variable, and the main dimensions of a PCA carried out on the sensory data, as stated by trained assessors, are the main (or unique) explanatory variables. Hence, this technique can be used to map consumer data onto the assessors’ space and obtain the sensory properties that influence consumer acceptability. Four alternative preference models exist to fit the data: vector, circular, elliptical, and quadratic, which were tested using the data of this study (Resano et al., 2010; Bonanya et al., 2014). In this study, external preference mapping was performed after consumer hedonic data were submitted to cluster analysis.

**Statistical Analysis**

The consumer test and the QDA means values were analyzed by one-way ANOVA and Tukey’s mean difference (P < 0.05) test, considering the sample as fixed effect (Granato et al., 2014). The QDA data were also subjected to PCA using the correlation matrix, whereas the hierarchical cluster analysis (HCA) was applied to the overall liking data to identify groups of consumers with similar preferences. In the HCA, Euclidean distances (dissimilarity), Ward’s method (agglomeration method), and automatic truncation were used (Cruz et al., 2013).

To identify attributes that were related positively with product acceptance, 2 methodologies were used: PLS and ANN, considering that these methods have opposite behavior with respect to linearity of the data. In both models, the overall liking was coded on a binomial scale (0/1), representing, respectively, rejection (value 0) for overall liking values among 1.00 to 5.99, and acceptance (value 1) for overall liking values from 6.00 to 9.00; Vidigal et al., 2011). To identify natural variation among consumers in the product acceptance test, the variables were not subject to any pre-processing of data. The PLS model quality was evaluated by 2 parameters: coefficient of determination (R²), related to the variation in the dependent variable explained by the regression model, and the cumulative index (Q²), which measures the global contribution of the first components as to predictive quality of the model (Donadini et al., 2012; Aquino et al., 2014).

The number of input neurons corresponds to the number of input variables into the neural network, and the number of output neurons is similar to the number of desired output variables. The input data set comprised the sensory attributes obtained from QDA. The output layer included 2 neurons corresponding to 2 possible acceptance classes, which can be classified according to overall liking (OL) as accepted (OL > 6.0) or rejected (OL ≤ 6.0). The target output was 1.0 in the correct class output, and 0.0 in the other. The outputs are probabilities of membership in every class, provided that the total across all output units is 1.0. The original data sets (280 samples) were randomly divided into a training set (75%) and a test set (25%) for all network topologies tested. The number of neurons in the hidden layer was obtained by trial and error. The stopping criterion was 1,000 epochs or maximum sum of square errors (SSE; Eq. [1]), equal to 0.001 during training. Various performance measures are computed during the test. Each predicted value was compared with the experimental value to test the network performance. For this purpose, mean square error (MSE) was calculated using Eq. [2]. On the lower values of MSE, the network predicts the values more truly, and the correlation coefficient (r; Eq. [3]) gives information about performance. A correlation coefficient close to 1 indicates that learning and prediction are successful:

\[
SSE = \sum_{i=1}^{N} (O_i - T_i)^2, \tag{1}
\]

\[
MSE = \frac{\sum_{i=1}^{N} (O_i - T_i)^2}{N}, \quad \tag{2}
\]

\[
r = 1 - \sqrt{\frac{\sum_{i=1}^{N} (O_i - T_i)^2}{\sum_{i=1}^{N} (O_i - T_m)^2}}, \tag{3}
\]

where \(O_i\) is the \(i\)th actual value, \(T_i\) is the \(i\)th predicted value, \(N\) is the number of data, and \(T_m\) (median value) is given by

\[
T_m = \frac{\sum_{i=1}^{N} O_i}{N}. \tag{4}
\]

A confusion matrix was performed to determine sensitivity (Se; true positive rate), specificity (Sp; true
negative rate), and accuracy (Acc) values to evaluate
the classification performance of the binary classifier
system (Xu et al., 2012). The Se, Sp, and Acc were
determined as follows and used to evaluate the perform-
ance of the pattern classification techniques for each
class (Xu et al., 2012):

\[
Se = \frac{TP}{TP + FN},
\]

\[
Sp = \frac{TN}{TN + FP},
\]

\[
Acc = \frac{TN + TP}{TN + TP + FN + FP},
\]

where TP, FN, TN, and FP denote the numbers of true
positives, false negatives, true negatives, and false posi-
tives, respectively. For calculation purposes, samples
classified as accepted were considered positive, and the
rejected samples were considered as negative. Further-
more, a sensitivity analysis was performed to provide a
measure of the relative importance among the inputs of
the neural network model and to illustrate the model
output variability as a function of the change in the
input variable.

The JAR scores were evaluated through penalty
analysis, being considered significant when more than
20% of consumers evaluated the sample above or below
JAR (Drake et al., 2011; Narayanan et al., 2014). The
purchase intent data, computed in binomial form (1,
representing buy, >3; 0, representing not buy, ≤3),
were correlated with data from the JAR scale and ac-
ceptance test by LR to find the affective characteristics
that most influenced the purchase intent of the product
(Cruz et al., 2011).

Except for ANN, all analyses were performed using
the software XLSTAT version 2013 (Addinsoft, Paris,
France). The ANN were modeled by the software Statis-
tica v. 8.0 (StatSoft Inc., Tulsa, OK) using a supervised
multilayer perceptron network to predict the sensory
acceptability of DL samples obtained using QDA (Cruz
et al., 2009, 2011).

RESULTS AND DISCUSSION

Sensory Profile

Table 3 shows the average attributes obtained in the
QDA of DL. The sensory profile was composed of 15
descriptors: apparent adhesiveness, apparent viscosity,
color, brightness, heated milk aroma, caramel aroma,
sweet taste, heated milk taste, caramel taste, butter,
pungent aftertaste, oil layer, adhesiveness, viscosity,
and sandiness. Significant differences were observed be-
tween all attributes for all samples (P < 0.05). In par-
cular, significant differences were observed for some
sensory descriptors such as brightness (ranging from
2.26 to 13.07), viscosity (ranging from 0.865 to 13.32),
and sandiness (ranging from 0.705 to 13.82), among
others. With respect to the samples, DL I, IV, and
VII showed extreme scores on the attributes bright-
ness, color, adhesiveness, and viscosity (apparent and
oral) and caramel and heated milk (aroma and taste),
whereas DL II presented significant scores for the attri-
bute sandiness. Finally, DL V and VI showed a higher
number of attributes with intermediate scores. Overall,
the results indicate wide variety in the products offered
to the consumer, resulting from different formulations
and parameters used in the manufacturing process, such
as intensity of heat treatment and amounts of sodium
bicarbonate, sucrose, and glucose, of which the latter is
a facultative ingredient.

The 2-dimensional PCA accounted for 87.2% (68.3%
and 18.9% for dimensions I and II, respectively, as shown
in Figure 1), and DL II was characterized by the attrib-
ute sandiness and, to a lesser extent, by the attributes
apparent adhesiveness, viscosity, and adhesiveness.
Sample IV was characterized by pungent aftertaste,
color, caramel taste and aroma, whereas samples V and
VI presented sweet taste, caramel taste and aroma, and
color. Samples I and III were characterized by buttery
taste and oil layer, whereas sample VII was charac-
terized by heated milk aroma and heated milk taste. As
can be seen in Figure 1, the first principal component
was characterized by the attributes brightness, pungent
aftertaste, butter taste, and viscosity, whereas the sec-
ond principal component was characterized by sandi-
ess, heated milk aroma, and color.

Consumer Test

The results of the consumer test are shown in Table
4. Corroborating the data from the sensory profile ob-
tained by QDA, significant differences were observed
for all attributes (P < 0.05), suggesting again het-
erogeneous products from the sensory point of view.
Samples VI, V, and II showed the highest values for
overall acceptance (7.5, 6.8, and 6.5, respectively; P
< 0.05), with similar findings for appearance (7.5, 6.6,
and 7.2, respectively; P < 0.05), aroma (6.8, 6.6, and
6.4, respectively; P < 0.05), taste (7.4, 7.0, and 6.3,
respectively; P < 0.05) and texture (7.6, 6.9, and 6.2,
respectively; P < 0.05). In contrast, samples I and VII
presented lower scores for overall appearance (4.6 and
4.1, respectively; P < 0.05), ranging from 4 (extremely
disliked) to 6 (liked slightly) for the other attributes,
Table 3. Average sensory scores for 7 dulce de leche samples

<table>
<thead>
<tr>
<th>Attribute</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>4.4d</td>
<td>5.7d</td>
<td>9.5c</td>
<td>14.5a</td>
<td>11.6b</td>
<td>12.4b</td>
<td>0.2e</td>
</tr>
<tr>
<td>Brightness</td>
<td>13.1a</td>
<td>9.6bc</td>
<td>11.1ab</td>
<td>2.3e</td>
<td>7.1d</td>
<td>8.0de</td>
<td>12.4a</td>
</tr>
<tr>
<td>Apparent adhesiveness</td>
<td>1.9f</td>
<td>12.1c</td>
<td>4.5f</td>
<td>13.6a</td>
<td>6.7d</td>
<td>6.1bc</td>
<td>4.3f</td>
</tr>
<tr>
<td>Apparent viscosity</td>
<td>1.3f</td>
<td>11.5bc</td>
<td>3.1f</td>
<td>13.6a</td>
<td>7.9f</td>
<td>5.7d</td>
<td>3.0f</td>
</tr>
<tr>
<td>Heated milk aroma</td>
<td>9.0f</td>
<td>8.8f</td>
<td>6.7f</td>
<td>2.9f</td>
<td>3.5f</td>
<td>4.0f</td>
<td>12.0f</td>
</tr>
<tr>
<td>Caramel aroma</td>
<td>5.6f</td>
<td>7.0f</td>
<td>8.5f</td>
<td>13.0f</td>
<td>10.9f</td>
<td>10.5f</td>
<td>2.4f</td>
</tr>
<tr>
<td>Sweet taste</td>
<td>11.1a</td>
<td>10.9a</td>
<td>10.5f</td>
<td>10.0ab</td>
<td>10.2f</td>
<td>10.5f</td>
<td>8.8b</td>
</tr>
<tr>
<td>Heated milk taste</td>
<td>9.0a</td>
<td>8.1b</td>
<td>7.3c</td>
<td>3.0f</td>
<td>4.6f</td>
<td>4.2f</td>
<td>12.1a</td>
</tr>
<tr>
<td>Caramel taste</td>
<td>6.8e</td>
<td>8.0d</td>
<td>9.5e</td>
<td>13.2f</td>
<td>11.1d</td>
<td>11.4d</td>
<td>2.8e</td>
</tr>
<tr>
<td>Butter taste</td>
<td>7.2b</td>
<td>2.4d</td>
<td>7.9b</td>
<td>1.4d</td>
<td>3.3d</td>
<td>4.7d</td>
<td>7.0b</td>
</tr>
<tr>
<td>Pungent aftertaste</td>
<td>5.3d</td>
<td>6.8d</td>
<td>5.2d</td>
<td>8.4d</td>
<td>6.3d</td>
<td>5.4d</td>
<td>2.9d</td>
</tr>
<tr>
<td>Oil layer</td>
<td>6.2d</td>
<td>2.5d</td>
<td>7.6d</td>
<td>2.0f</td>
<td>3.4d</td>
<td>4.7d</td>
<td>5.4d</td>
</tr>
<tr>
<td>Adhesiveness</td>
<td>0.7c</td>
<td>10.5b</td>
<td>2.4e</td>
<td>13.3f</td>
<td>6.5e</td>
<td>5.4e</td>
<td>2.9d</td>
</tr>
<tr>
<td>Viscosity</td>
<td>0.9c</td>
<td>10.4b</td>
<td>2.2c</td>
<td>13.5f</td>
<td>6.8e</td>
<td>5.0f</td>
<td>2.6c</td>
</tr>
<tr>
<td>Sandiness</td>
<td>0.7c</td>
<td>13.8e</td>
<td>0.8c</td>
<td>0.2c</td>
<td>0.2c</td>
<td>0.2c</td>
<td>0.2c</td>
</tr>
</tbody>
</table>

\[a-f^*\] Means with the different letters in the same rows are significantly different \((P < 0.05)\) using Tukey’s test.

Figure 1. Principal component analysis of dulce de leche samples, showing factors (F)1 and F2, of attributes and samples (brands I, II, III, IV, V, VI, VII), as given in Table 3. Color version available online.
with no attribute scoring above 6, which represents the lower limit of acceptance zone on a 9-point scale (Vidigal et al., 2011).

Interestingly, samples II and VI showed higher scores for appearance compared with sample VII, which had the lightest color according to the trained panel, showing the importance of color in product appearance. The higher scores found for sample VI in overall impression may be related to descriptive characteristics that, in general, did not present discrepant mean values, given the expectation of a greater number of consumers. It is important to emphasize that the composition of sample VI according to the label was simple: besides milk and sucrose, the formulation contained only sodium bicarbonate and potassium sorbate. This indicates the possibility of manufacturing DL with better acceptance and simplified composition, and reducing manufacturing costs by using a less complex and more economical technology. Despite it is not being within the values recommended for DL, the attribute sandiness did not influence overall impression, because DL II was accepted despite showing significant scores for this attribute.

Cluster Analysis

Cluster analysis was performed with the aim of targeting consumers based on their global acceptability. Three clusters were identified with 74, 24, and 47 consumers each, corresponding to 59.2, 19.2, and 37.6% respectively (Table 5). Because 88 of the 125 consumers were women (70.4%), all clusters had a preponderance of women. In respect to age, because the study was developed in a university center, we observed a preponderance of younger adults, aged 18 to 24 yr in clusters 1 and 3 (68.5 and 59.6% of consumers, respectively), whereas adults aged 25 to 34 yr predominated in cluster 2 (corresponding to 58.3% of consumers in this cluster). Regardless of the segment formed, samples II, V, and VI showed the highest scores for overall acceptability (values between 7.1 and 8.1). Analyzing the differences between clusters for each food, we can observe that members of cluster 1 had a greater tendency to accept DL samples, regardless of the heterogeneity of the DL samples. Members of cluster 1 gave the highest acceptance scores for all 7 DL, with statistical difference \( P < 0.05 \) compared with the other 2 clusters, except for cluster 2 for sample V and cluster 3 in sample II. The relevance of cluster 1 is its high concentration of consumers. Additionally, the high score of sample II in cluster 3 suggests that sandiness is not a relevant parameter for rejection of that product. This reinforces the findings of previous studies in which it was observed that sandiness is tolerated in different ways with respect to consumers of fresh milk (Giménez et al., 2008).

Table 4. Overall liking attribute means from consumer acceptance testing of dulce de leche

<table>
<thead>
<tr>
<th>Item</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appearance liking</td>
<td>4.7(^c)</td>
<td>7.2(^a)</td>
<td>5.1(^c)</td>
<td>4.8(^c)</td>
<td>6.6(^b)</td>
<td>7.5(^a)</td>
<td>3.7(^d)</td>
</tr>
<tr>
<td>Aroma liking</td>
<td>5.2(^c)</td>
<td>6.4(^b)</td>
<td>6.0(^c)</td>
<td>5.8(^c)</td>
<td>6.6(^b)</td>
<td>6.8(^a)</td>
<td>4.6(^b)</td>
</tr>
<tr>
<td>Taste liking</td>
<td>4.6(^c)</td>
<td>6.3(^b)</td>
<td>6.1(^c)</td>
<td>5.0(^c)</td>
<td>7.0(^b)</td>
<td>7.4(^a)</td>
<td>4.0(^b)</td>
</tr>
<tr>
<td>Texture liking</td>
<td>4.6(^c)</td>
<td>6.2(^b)</td>
<td>5.5(^b)</td>
<td>5.0(^b)</td>
<td>6.9(^b)</td>
<td>7.6(^b)</td>
<td>5.6(^d)</td>
</tr>
<tr>
<td>Overall liking</td>
<td>4.6(^c)</td>
<td>6.5(^a)</td>
<td>5.8(^b)</td>
<td>5.0(^b)</td>
<td>6.8(^b)</td>
<td>7.5(^a)</td>
<td>4.1(^b)</td>
</tr>
<tr>
<td>Purchase intent</td>
<td>2.2(^a)</td>
<td>3.3(^b)</td>
<td>2.9(^d)</td>
<td>2.4(^b)</td>
<td>3.7(^b)</td>
<td>4.1(^b)</td>
<td>1.9(^b)</td>
</tr>
</tbody>
</table>

\(^{a-f}\)Means with the different letters in the same rows are significantly different \( P < 0.05 \) using Tukey’s test.

\(^{1}\)Data represent 125 consumers. Liking attributes were scored on a 9-point hedonic scale, where 1 = dislike extremely and 9 = like extremely; purchase intent was scored on a 5-point hedonic scale, where 1 = “I definitely would not buy” and 5 = “I definitely would buy”; just-about-right (JAR) scale was scored on a 9-point scale where 1 to 4 = below JAR, 5 = JAR, and 6 to 9 = above JAR. The JAR results indicate the percentage of assessor that selected these options; the number in the parentheses is the mean decrease calculated when the percentage of citations exceeded 20%.
Table 5. Cluster analysis results of overall liking of dulce de leche by consumers.

<table>
<thead>
<tr>
<th>Brand</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>5.9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.0&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.9&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>II</td>
<td>6.9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.9&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.2&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>III</td>
<td>6.4&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.5&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>IV</td>
<td>5.7&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4.0&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.7&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>V</td>
<td>7.2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6.0&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>VI</td>
<td>8.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.0&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>VII</td>
<td>5.3&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.0&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.1&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a,b</sup>Means with different letters in the same rows are significantly different (P < 0.05) using Tukey’s test.

<sup>1</sup>Data represent 125 consumers reflecting their sex, age, and liking. Liking attributes in clusters 1, 2, and 3 were scored on a 9-point hedonic scale, where 1 = dislike extremely and 9 = like extremely.

**JAR Scale and Penalty Analysis**

Table 4 shows the penalty analysis performed using data from the JAR test for the attributes color, sweet taste, caramel taste, and consistency. Sample VI showed small mean decreases in all attributes evaluated (0.4, 0.5, 0.7, and 0.8, respectively), confirming the high acceptability demonstrated by overall impression, and suggesting an optimized formulation from the sensory point of view. The DL VII showed the highest response percentages in color below JAR (94.4%), and greater mean drop value used to estimate the penalty due to its below JAR (3.1), which were the highest among all the samples. This corroborates the negative acceptance in overall impression and shows that the attribute color is important as a criterion for sensory quality in the eyes of the consumer. The opposite behavior produced the same effect, as the excessive browning of DL IV showed high frequency (89.6%) and mean drop (2.3), considering it much too brown (above JAR). The higher frequency of caramel taste below JAR was found for the sample VII (82.4%), which, as occurred earlier for JAR color attribute, presented a high mean drop (2.3), indicating the strong influence of this attribute on its acceptance and a direct relationship with color due to excessive Maillard reaction. In general, DL samples I, IV, and VII presented the lowest mean values for overall liking. The mean drops indicate that this may be due to the sweet taste below JAR and caramel taste below JAR for DL I, brownish above JAR and consistency above JAR for the DL IV, and brownish below JAR and caramel taste below JAR for DL VII.

According to Villegas et al. (2010), a specific attribute is present in optimal levels in a product when a at least 70% of the answers are in the JAR group. Based on this, the response frequencies showed that DL VI can be considered JAR with respect to its consistency, because 74.4% of consumers rated the sample JAR for consistency, although a significant number of responses was also observed for the attribute caramel taste, with 61.6%.

**LR and Purchase Intent**

Table 6 shows the attributes that influenced purchase intent of DL using LR. One predictive value (87.54%) equivalent to a good ability to fit the experimental data was observed. The results suggest that buying products is a multidimensional parameter, influenced mainly by 2 attributes: overall liking and taste, with odds ratio values of 3.23 and 1.81, respectively. The higher scores for overall liking and taste suggest that these attributes are more critical for the purchase of the product, with a purchase probability being 3.23 and 1.81 times higher (compared with not being purchased; P < 0.0001) for every 1-unit increase in overall liking score (based on a 9-point hedonic scale). However, texture is also relevant, with a purchase probability being 1.21 times higher compared with not being purchased (P < 0.001), resulting in 1-unit increase of the overall liking score.

Although the LR and penalty analysis findings were useful, they must be interpreted with care because most of the consumers who took part in the study were students at the university. Further study should be performed using a general population to confirm these results.

**Drivers of Liking Using PLS and ANN**

The use of PLS generated 3 components, with Q<sup>2</sup> acum (a measure of cross-validation of the PLS model), R<sup>2</sup> X acum (percentage variability explained in 3 dimensions of the dependent variable), and R<sup>2</sup> Y acum (percentage variability explained in 3 dimensions of the independent variable) values of 0.564, 0.719, and 0.411, respectively. Thus, these values indicate low predictive ability of the model and little explanation of the variability of the data, which may be related to the heterogeneity of formulations, which contributed to the nonlinear behavior data, once samples with high scores in certain descriptors presented opposite behavior for other descriptors. In this context, the PLS does not seem adequate to estimate drivers of liking of DL; therefore, a nonlinear method, such as ANN, was needed.

The performance of ANN was evaluated using prediction accuracy and MSE of training and testing sets. The proximity of the Acc values to 100% and closeness of the MSE values to 0 show the efficiency of the prediction of ANN model. The best ANN architecture was composed of 15 input neurons, 5 hidden neurons, and 2 output neurons using logistic and hyperbolic tangent function in hidden and output layers, respectively.
The receiver operating characteristic (ROC) curve is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve analysis demonstrated Se (true positive rate), Sp (true negative rate), and Acc values of 98.0, 100.0, and 98.6% for the whole data set. The high accuracy of the ANN model can be corroborated by the low MSE value of 1.73%, with 100% of recognition and prediction ability for the accepted class, whereas the rejected class presented 95 and 93% recognition and prediction ability for the training and testing data sets, respectively. The sensitivity analysis showed color, apparent viscosity, adhesiveness, and caramel taste as the sensory attributes with the most discriminative power, presenting variable of importance in projection coefficients of 7.009, 3.915, 3.501, and 2.321 for the whole data set, respectively. In this context, improving the acceptance of DL requires a reformulation of the product and process parameters that influence these attributes, such as control of temperature and time of heat treatment and addition of ingredients such as sucrose, glucose, and sodium bicarbonate.

**External Preference Mapping**

Figure 2 shows the external preference mapping obtained from the regression of the acceptance data of each cluster on the 2-dimensional PCA. For cluster 1, a circular model was obtained from the existence of an ideal point (−0.723, 0.542), whereas quadratic and elliptical models were obtained for cluster 2 and cluster 3, respectively, with the existence of saddle points that are shown on the map. For a specific individual, the ideal point shows the location of his or her most preferred sample (a global maximum in terms of preference), whereas the saddle point represents a threshold where the variability of the preference is low, before increasing or decreasing in the opposite direction (Resano et al., 2010).

As shown in the 2-dimensional map, samples V and VI were located in a region that included acceptance values above the overall value for 80 to 100% of consumers. The ideal point exists in this region, and these samples are associated with the attributes color, sweet taste, caramel taste, and caramel aroma. The saddle point region precedes the region in which the acceptance values were 60 to 80% above the overall acceptance, where samples II and III are located. Sample II was associated with sandiness to a greater extent, followed by the attributes apparent and oral adhesiveness, and apparent and oral viscosity, whereas sample III was related to oil layer and butter taste.

**Sensory Profiling of the Ideal Products**

Besides establishing the sensory profile and the characteristics that drive product acceptance by consumers, there is a need to determine the sensory characteristics that should be exhibited by the ideal product. This procedure is performed by reverse regression, which, in our case, was a PLS regression of the sensory scores on either the preference space coordinates (obtained from circular model, cluster 1) of the ideal product in their respective spaces.

We determined that the optimum DL was characterized by high scores for the attributes sweet taste, caramel taste, brightness, color, and caramel aroma (10.3, 9.1, 9.0, 8.5, and 8.4, respectively), intermediate scores for apparent adhesiveness, heated milk taste, apparent viscosity, heated milk aroma, and adhesiveness (7.1, 6.8, 6.7, 6.6, and 6.0, respectively), and low scores for viscosity, pungent aftertaste, butter taste, oil layer, and sandiness (6.0, 5.8, 4.8, 4.6 and 2.2 respectively), considering a 15-cm linear scale.

It is important to note that the sensory attributes having high values were in complete agreement with the preference mapping findings, whereas for the ANN results, the attributes adhesiveness and viscosity were

---

**Table 6. Parameter estimates, probability, and odds ratio estimates for purchase intent of dulce de leche**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter estimate</th>
<th>P &gt; χ²</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appearance</td>
<td>0.18</td>
<td>0.052</td>
<td>1.19</td>
</tr>
<tr>
<td>Aroma</td>
<td>−0.01</td>
<td>0.907</td>
<td>0.99</td>
</tr>
<tr>
<td>Taste</td>
<td>0.59</td>
<td>&lt;0.0001</td>
<td>1.81</td>
</tr>
<tr>
<td>Texture</td>
<td>0.19</td>
<td>0.034</td>
<td>1.21</td>
</tr>
<tr>
<td>Overall liking</td>
<td>1.17</td>
<td>&lt;0.0001</td>
<td>3.23</td>
</tr>
<tr>
<td>Color (JAR²)</td>
<td>0.11</td>
<td>0.257</td>
<td>1.12</td>
</tr>
<tr>
<td>Sweet taste (JAR)</td>
<td>0.12</td>
<td>0.267</td>
<td>1.13</td>
</tr>
<tr>
<td>Caramel taste (JAR)</td>
<td>−0.02</td>
<td>0.846</td>
<td>0.98</td>
</tr>
<tr>
<td>Consistency (JAR)</td>
<td>−0.05</td>
<td>0.611</td>
<td>0.95</td>
</tr>
</tbody>
</table>

**Footnotes:**
1. Based on the logistic regression analysis, using a full model with 9 sensory attributes. The analysis of maximum likelihood estimates was used to obtain parameter estimates. Significance of parameter estimates was based on the Wald χ² value at P < 0.05.
2. "Just-about-right scale."
not noted. This is understandable because the preference mapping was obtained by a previous segmentation of consumers in relation to the overall acceptance scores, whereas, for ANN, the variability of the consumers was maintained. Overall, it is more prudent and adequate to look closely at attributes with high values in the reverse regression, because their improvement could positively affect overall liking.

From an industrial perspective, intermediate heat treatments may be used, because attributes such as caramel color, aroma, and taste are influenced by this operational parameter, together with the formulation ingredients, such as sucrose, glucose (which contributes to the brightness and browning), and sodium bicarbonate (acidity regulator, which positively influences the Maillard reaction).

CONCLUSIONS

Expansion into new consumer markets requires standardization of sensory characteristics of processed products, resulting in greater consumer acceptance. For DL in the present study, large heterogeneity was observed from the sensory point of view, and the products most appreciated by consumers had intermediate scores for the sensory attributes established in the descriptive test; this trend was observed even after consumer segmentation. However, the need for effective optimization remains, because no product was considered ideal for the attributes under study. In general, the sensometric methodologies were useful and indicated that sensory optimization of DL is a multidimensional process, requiring simultaneous adjustments in the attributes appearance, aroma, taste, and texture. From a practical standpoint, heat treatment and quantity of ingredients used in the formulation should be standardized. The ideal DL from a consumer’s perspective was characterized by high scores for the attributes sweet taste, caramel taste, brightness, color, and caramel aroma. These results may be useful to dairy producers in the manufacturer of DL with optimized sensory quality and high probability of acceptance by consumers, thus resulting in increased revenues for the producer.

REFERENCES


Aquino, L. F. M. C., A. C. O. Silva, M. Q. Freitas, T. L. Felicio, A. G. Cruz, and C. A. Conte-Junior. 2014. Identifying cheese whey...


