ABSTRACT

Developing innovative products that satisfy various groups of consumers helps a company maintain a leading market share. The hedonic scale and just-about-right (JAR) scale are 2 popular methods for hedonic assessment and product diagnostics. In this paper, we chose to study flavored liquid milk because it is one of the most necessary nutrient sources in China. The hedonic scale and JAR scale methods were combined to provide directional information for flavored liquid milk optimization. Two methods of analysis (penalty analysis and partial least squares regression on dummy variables) were used and the results were compared. This paper had 2 aims: (1) to investigate consumer preferences of basic flavor attributes of milk from various cities in China; and (2) to determine the improvement direction for specific products and the ideal overall liking for consumers in various cities. The results showed that consumers in China have local-specific requirements for characteristics of flavored liquid milk. Furthermore, we provide a consumer-oriented product design method to improve sensory quality according to the preference of particular consumers.

Key words: partial least squares regression, penalty analysis, just-about-right scale, local consumer preference difference

INTRODUCTION

New product introductions are critical to the growth and continuing success of a company. Developing innovative products that satisfy various groups of consumers helps maintain a company’s market share. Consumer judgments are needed to exploit new markets based on preference understanding (Ruan and Zeng, 2004). Designing products that satisfy consumers’ flavor preferences is the core consideration in product development. Sensory quality is one of the most important factors affecting the final choices of consumers. Qualitative and quantitative studies help to identify the factors that drive consumers’ liking, providing the understanding needed to optimize new products and enhance the profitability of existing products (Raz et al., 2008).

In market research, the hedonic scale and the just-about-right (JAR) scale are 2 popular methods for hedonic assessment and product diagnostics. The hedonic scale is a balanced bipolar scale centered around neutral, with categories labeled with phrases representing various degrees of liking (Villanueva and Da Silva, 2009; Lim and Fujimaru, 2010; Lim, 2011). The JAR method is a direct way to solicit feedback from a consumer, asking whether a product is just right or has too much or too little of a certain characteristic (Popper and Kroll, 2005; Rothman, 2007; Plaehn and Horne, 2008). Usually, the hedonic and JAR scales are combined to provide directional information for product reformulation. Penalty analysis (PA) is one of the most commonly used analysis methods for the JAR scale, and it has been extensively used in the food industry (Meullenet et al., 2007; Paczkowski, 2009; Xiong and Meullenet, 2009). Penalty analysis could identify potential directions for product improvement. It assists in identifying attributes that cause an increase or decrease in hedonic scale associated with sensory attributes not at optimal levels in a product, allowing the product developer to decide what sensory properties should be improved or adjusted (Paczkowski, 2009). Regression-based method is well applied for analyzing JAR scale and hedonic scale of overall liking score (Xiong and Meullenet, 2004; Plaehn and Horne, 2008). Partial least squares (PLS) regression on dummy variables (Xiong and Meullenet, 2006; Worch et al., 2010) can be used to analyze JAR data, which have better prediction properties for overall liking than original variables. Both penalty analysis and regression-based methods could estimate the potential improvement margin when the attributes are not at the just-about-right level.
With development of a national economy and the improvement of living standards in China, flavored liquid milk has become an essential daily nutrition source of Chinese residents. Today, competition in the food industry is intense, and flavored liquid milk is expected to develop toward the diverse requirements of consumers, to adapt to the different consumption levels, preferences, and nutritional requirements of consumers.

In this study, we analyzed local preference characteristics for flavored liquid milk for residents in various areas of China. An acceptance test was adopted to investigate overall liking (hedonic scale), and the perceptive intensity of attributes (JAR scale) of formulated milk product. Two methods of analysis—PA and PLS regression on dummy variables—were used to identify the improvement margin with the attributes that are not at the JAR level. We focused on the following 2 aims: (1) to investigate the consumers' preferences of milk products and their flavor attributes in various cities in China; and (2) to determine the improvement direction for specific products and the ideal overall liking for various cities. Furthermore, we provide a consumer-oriented product design method to improve sensory quality according to the preference of particular consumers.

**MATERIALS AND METHODS**

**Sample Preparation**

Flavored liquid milks were designed with different flavors referring to Chinese National Standards (2007, 2008). The main components include milk, sugar, lactic acid, and citric acid, and thickener (carboxymethylcellulose, CMC). Milk was used as the raw material, protein content was not less than 1.0%, and sugar and acid (lactic acid and citric acid) were added to make the liquid milk have a certain flavor. Thickener was added for stability and smooth mouthfeel. An orthogonal test design was made with these 4 factors with 4 levels each (see Table 1 for details).

<table>
<thead>
<tr>
<th>Component</th>
<th>Level</th>
<th>Milk (% protein)</th>
<th>Sugar (g)</th>
<th>Acid (g)</th>
<th>CMC (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0</td>
<td>7.0</td>
<td>0.40</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.1</td>
<td>8.0</td>
<td>0.50</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.2</td>
<td>9.0</td>
<td>0.60</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.3</td>
<td>10.0</td>
<td>0.70</td>
<td>0.45</td>
<td></td>
</tr>
</tbody>
</table>

1CMC = carboxymethylcellulose (thickener).

Sixteen types of liquid milk were formulated based on the orthogonal test design. According to the ISO standard (ISO, 1993), assessors were selected and trained to meet the sensitivity requirement, and the sensory panel consisted of 20 assessors who were recruited from the sensory panel trained and tested regularly by the sensory analysis laboratory, and 12 of which were male. Most of them were familiar with flavored milk and had an average age of 35 yr, ranging from 28 to 39 yr.

The conditions of the sensory evaluation environment met the international standard (ISO, 2007). The room temperature was air-conditioned to maintain a temperature of 18°C. All 16 samples were presented monadically in random order with samples identified by 3-digit random code on the cup. To prevent sensory and mental fatigue resulting from tasting all samples, the 16 samples were divided into 4 groups and provided to assessors as 4 samples per session. During the sensory evaluation procedure, plain bread and water were provided to panelists to cleanse the palate. Assessors were asked to rest for 3 min between sessions. The 16 samples were individually assessed for sensory attributes of appearance, texture, aroma, and flavor. Descriptors were identified and selected for establishing the sensory profile.

A vocabulary of appearance, texture, aroma, and flavor was developed following the ISO standard (ISO, 1994). Finally, 12 sensory descriptors were selected by using a multidimensional approach, and the intensities of the selected attributes were evaluated.

The 12 sensory attributes (including appearance, texture, aroma, and flavor) were scored on a 15-cm line scale that had numerical anchors (0 and 15) at both ends. The numerical anchor 0 = none, 15 = extremely strong, and 0 to 15 means the trend of low intensity to high intensity. The average results of the sensory assessors were calculated for further processing.

**Investigation of Consumer Perception**

Three areas of China (north, east, and southwest) were chosen to represent typical flavor preferences, and Beijing, Shanghai, and Chengdu were selected in these 3 areas, respectively. Six hundred fifty-six adult consumers (16 to 45 yr of age) participated in this study (218 in Beijing, 218 in Chengdu, and 220 in Shanghai). The ratio between male and female was almost 1:1, and the distribution of age was 5:3:2 for 16 to 25 yr, 26 to 35 yr, and 36 to 45 yr. Respondents were recruited to a central location to participate in this study. Local flavor preferences for consumers from various cities are very helpful to improve the sensory quality of flavored milk. In particular, local citizens who live in a city for a
long time represent the typical flavor preference of the city. Individuals who consume flavored milk frequently are more important for flavor design. Therefore, local citizens who had lived in the city for more than 2 yr and consumed flavored milk at least 3 times per week were chosen to participate.

Five key sensory attributes were evaluated by the selected consumers: sourness, sweetness, fresh milk flavor, thickness, and smoothness. Samples were evaluated for consumer preference using a 5-point hedonic scale for liking of flavored liquid milk (1 = dislike very much, 2 = dislike slightly, 3 = neither like nor dislike, 4 = like slightly, 5 = like very much) and for intensities of the key sensory attributes by JAR scale (1 = much too little, 2 = too little, 3 = just about right, 4 = too much, 5 = much too much). In the consumer test, the room temperature was air-conditioned to maintain a temperature of 18°C, and the milk samples were kept at the same temperature. All samples were presented in random order at the same time.

**Penalty Analysis**

Penalty (or mean-drop) analysis is a method used extensively in sensory data analysis to identify potential directions for the improvement of products. It determines what attributes can lead to an increase in overall liking (Paczkowski, 2009) by showing how many points you lose for having a product attribute be “too much” or “too little” for a consumer. Penalty analysis can aid product developers in understanding differences in product preference. It can also help in understanding the basis for consumer segmentation and product substitutability (Rothman, 2007).

Penalty analysis itself is based on the idea that the maximum hedonic score or overall liking will occur at the JAR point (Plaehn and Horne, 2008). Thus, PA used the data collected on the JAR scale and the liking scores on hedonic scale.

The PA consisted of 3 steps: (1) collapse the 5-point JAR scale into 3 categories (too little, JAR, and too much) for each attribute; (2) calculate the mean overall liking score from the hedonic scale and the percentage of respondents represented in each of the 3 categories; (3) calculate mean drops by subtracting the mean overall liking score for the JAR group from the mean overall liking score of the “too much” or “too little” categories (Anon, 2003; Rothman, 2007).

**PLS on Dummy Variables**

Partial least squares regression was used to find relationships between independent variables and dependent variables (numerical variables); PLS is a statistical method that bears some relation to principal components regression. However, instead of finding hyperplanes of minimum variance between the response and independent variables, PLS finds a linear regression model by projecting the predicted variables and the observable variables to a new space.

Categorical variables can be used in a PLS model if they are converted to dummy variables. The relationship between overall liking of consumer preference and perceptive acceptance for the flavored liquid milk samples was established by PLS on dummy variables. The aim of PLS on dummy variables is to estimate the possible gain in liking if that attribute was at its ideal level.

Xiong and Meullenet (2006) proposed the use of PLS on dummy variables to analyze JAR data. There are 3 main steps for PLS on dummy variables:

- Transform each of the original JAR-scale variables into a pair of variables. The JAR scale has the “ideal” score as its midpoint, and it can be converted to dummy variables by calculating the difference between the real JAR score and ideal JAR score, that is: $Z_{11} = Z_{real} - Z_{ideal}$. Therefore, the original JAR scale is divided into 2 columns. The values of “too little” are negative (−) and the values of “too much” are positive (+).
- Establish the relationship between overall liking score and dummy variables by PLS. Dummy variables can be used with any regression procedure, and PLS regression was used in our research. Similar to PA, attributes that were not rated to be JAR with 20% or more consumers are selected to the regression model (Estiaga, 2011). The coefficients that affect liking are extracted. The regression weights help explain why liking is not optimal for the product and indicate the direction of improvement for the attributes.
- Calculate the mean drop of dummy variables for each attribute. Mean drop values can also be calculated by PLS on dummy variables, indicating the potential improvement margin if the attributes are improved to the JAR level. The mean drop associated with each dummy variable was calculated by multiplying dummy variables by their associated regression weights.

Analysis of variance and clustering analysis were conducted using SPSS 18.0 (SPSS Inc./IBM Corp., Armonk, NY) and PLS regression was conducted using Matlab 2015b (The MathWorks Inc., Natick, MA).
RESULTS AND DISCUSSION

Liquid Milk Sample Sensory Evaluation

Clustering analysis was used to visualize the similarity of 16 samples according to the intensity of sensory attributes. Samples that had unusual sensory qualities from technological production (e.g., precipitation and hierarchy) were abandoned, and samples clustered in one group (with similar sensory characteristics) are represented by a typical sample. Finally, 5 types of liquid milk with different flavors (samples A to E) were chosen for the consumer test; the composition of these samples is compared in Figure 1.

These 5 samples were evaluated by the sensory panel for 12 sensory attributes (including appearance, texture, aroma, and flavor), and a spider plot was created to provide a graphic representation of the quantitative descriptive analysis (Figure 2). Furthermore, ANOVA was conducted and significant differences were identified for the 5 samples ($F_{4,55} = 6.62; P \leq 0.05$). The ANOVA for difference analysis between samples with certain attributes showed that these 5 sensory attributes were more important, so they were selected. Therefore, 5 key sensory attributes, including 3 flavor attributes (sourness, sweetness, and fresh milk flavor) and 2 texture attributes (thickness and smoothness) were determined for the consumer test.

Consumer Preference Results

Total Preference. Figure 3 shows the results of average liking evaluation by hedonic scale of 5 flavored liquid milks by consumers from Beijing, Chengdu, and Shanghai. The average overall liking degree was much higher for samples C, D, and E than for samples A and B. From the perspective of differences among regions, samples C, D, and E were liked by consumers from all 3 cities, so the regional difference was limited; for samples A and B, the liking degree of consumers from Shanghai was higher than that of consumers from the other 2 cities. Furthermore, an ANOVA was conducted for 5 samples from 3 cities to identify significant differences for preferences of the 5 samples, and significant differences were identified among the samples ($F_{4,10} = 17.47; P \leq 0.05$). The hedonic liking score of sample B was lower than that of other samples, and sample E received the highest hedonic liking score.

JAR for Sensory Attributes. Results of the JAR evaluation of 5 flavored liquid milks by consumers from Beijing, Chengdu, and Shanghai are shown in Figure 4. More respondents from the 3 cities rated samples C, D,
and E as being at the just-about-right level for sourness, sweetness, fresh milk flavor, and thickness compared with samples A and B. The respondents from Beijing evaluated sample C as being at the just-about-right level for sourness (65.1%), sweetness (61.0%), fresh milk flavor (63.8%), and thickness (58.3%); sample D as being at the just-about-right level for sourness (56.4%), sweetness (60.6%), fresh milk flavor (61.0%), and thickness (59.2%); and sample E as being at the just-about-right level for sourness (68.3%), sweetness (61.0%), fresh milk flavor (63.8%), and thickness (58.3%).

Figure 2. Spider plot for sensory scores of flavored liquid milk samples. Color version available online.

Figure 3. Hedonic evaluation for Beijing, Chengdu, and Shanghai. A 5-point hedonic scale was used for liking of flavored liquid milk (1 = dislike very much, 2 = dislike slightly, 3 = neither like nor dislike, 4 = like slightly, 5 = like very much).
(64.2%), fresh milk flavor (63.3%), and thickness (66.5%). Therefore, consumers in Beijing considered the key attributes of sample E to be at the ideal level.

Respondents from Chengdu evaluated sample C as being at the just-about-right level for sourness (67.4%), sweetness (65.1%), fresh milk flavor (57.8%), and thickness (54.6%); sample D as being at the just-about-right level for sourness (63.8%), sweetness (59.6%), fresh milk flavor (54.6%), thickness (59.6%), and smoothness (51.4%); and sample E as being on the just-about-right level for sourness (67.9%), sweetness (62.8%), fresh milk flavor (51.4%), and thickness (57.3%). Therefore, the consumers in Chengdu considered the key attributes of sample C and E to be at the ideal level.

Respondents from Shanghai evaluated sample C as being at the just-about-right level for sourness (62.7%), sweetness (63.2%), fresh milk flavor (64.5%), and thickness (72.7%); sample D as being at the just-about-right level for sourness (59.5%), sweetness (63.2%), fresh milk flavor (60.9%), and thickness (68.2%); and sample E as being at the just-about-right level for sourness (61.8%), sweetness (63.6%), fresh milk flavor (53.6%), and thickness (59.1%). Therefore, the consumers in Shanghai considered the key attributes of sample C to be at the ideal level.

Generally, at least 70% of responses should be at the just-about-right level to conclude that a specific attribute is at its optimal level (Xiong and Meullenet, 2004). Therefore, only the thickness of product C was at its optimal level for Shanghai respondents.

From the JAR results of 3 cities, samples C and E showed a greater number of respondents for just-about-right level. In particular, consumers in Beijing considered the attributes of sourness, sweetness, fresh milk flavor and thickness of sample E to be at the just-about-right level; consumers in Chengdu considered the
key attributes of sample C and E to be the just-about-right level; and consumers of Shanghai assessed the key attributes of sample C to be at the just-about-right level. For samples A and B, sourness was too much, and sweetness, fresh milk flavor, and thickness were too little. These results are consistent with the sensory analysis for 5 samples.

Improvement of Sensory Quality of Flavored Liquid Milk

Results of PA. Figure 5 is an example of a PA plot for sample A of Beijing. The x-axis denotes the percentage of respondents in a non-JAR category and the y-axis represents the mean drops. Attributes located in the upper right quadrant are those needing to be improved (Meullenet et al., 2007). Because the typical skew cut-off percentage used in the industry is 20%, mean drops are not considered if the proportion of respondents who rate a certain attribute is <20%. Responses below 20% are too small to be considered and might not be sufficiently reliable (Meullenet et al., 2007; Rothman, 2007).

According to the penalty analysis plot, we summarize the attributes need to be improved mostly for the 3 cities in Table 2. The attributes were selected according to the boundary (20%; i.e., attributes that were not rated to be JAR with more than 20% consumers). For Beijing, the attributes that required most improvement were (where + indicates “too much” and − indicates “too little”) for sample A sourness (+), fresh milk flavor (−), sweetness (−), thickness (−), and smoothness (−); for sample B sourness (+), sweetness (−), fresh milk flavor (−), thickness (−), and smoothness (+); for sample C smoothness (+), sweetness (+), and thickness (−); for sample D smoothness (+), sourness (+), and thickness (−); and for sample E smoothness (+) and sweetness (+).

For Chengdu, the attributes that required most improvement for sample A were sourness (+), sweetness (−), fresh milk flavor (−), thickness (−), smoothness (−), for sample B were sourness (+), sweetness (−), fresh milk flavor (−), thickness (−), smoothness (−), for sample C were fresh milk flavor (−), thickness (−), smoothness (+), for sample D were sweetness (−), fresh milk flavor (−), thickness (−), and smoothness (+); and for sample E were sourness (−), fresh milk flavor (−), thickness (−), and smoothness (+).

For Shanghai, the attributes that required most improvement for sample A were sourness (+), sweetness (−), fresh milk flavor (−), thickness (−), and smoothness (+); for sample B were sourness (+), sweetness (−), fresh milk flavor (−), thickness (−), and smoothness (+); for sample C were sourness (−), fresh milk flavor (−), and smoothness (+); for sample D were sweetness (−), fresh milk flavor (−), thickness (−), and smoothness (+); and for sample E were sourness (−), fresh milk flavor (−), thickness (−), and smoothness (+).

Figure 5. Penalty analysis plot for sample A (Beijing). The dashed line represents the boundary of 20% consumers. Mean drop: the difference between the mean overall liking score for just-about-right group and mean overall liking score of the “too much” or “too little” categories. Color version available online.
smoothness (+); for sample D were sourness (+), and smoothness (+); and for sample E were sourness (−), fresh milk flavor (−), thickness (−), and smoothness (+).

**Results of PLS.** As noted earlier, PA usually uses 20% as a cutoff point to preselect important attributes. We also used a criterion of 20% for PLS on dummy variables. The statistics for PLS on dummy variables are shown in Table 3. The regression intercept indicates the potential maximum mean score when all the JAR-scale attributes are at the just-about-right level. The difference between the intercept and the predicted overall liking mean score can be interpreted as the overall mean drop. The results indicate that the PLS model predicted the observed overall liking mean scores for the products very well, and most of the predicted mean scores were identical or very close to the observed mean score. The overall mean drop indicates the maximum potential improvement margin on overall liking if the attributes that are not at an optimal level are modified to be just about right. For example, for sample B in Chengdu, the overall mean drop was 1.26, which was much higher than other samples, meaning that the maximum potential improvement margin on overall liking was 1.26.

Figures 6, 7, and 8 show the mean drop results from PLS on dummy variables method for Beijing, Chengdu, and Shanghai, respectively. Figure 6 indicates the PLS

---

**Table 2.** Comparison of attributes that most need improvement (percentage of respondents)\(^1\)

<table>
<thead>
<tr>
<th>Location and sample</th>
<th>Soursness</th>
<th>Sweetness</th>
<th>Fresh milk flavor</th>
<th>Thickness</th>
<th>Smoothness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TL</td>
<td>TM</td>
<td>TL</td>
<td>TM</td>
<td>TL</td>
</tr>
<tr>
<td>Beijing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>64.7</td>
<td>58.7</td>
<td>60.1</td>
<td>54.1</td>
<td>24.3</td>
</tr>
<tr>
<td>B</td>
<td>73.9</td>
<td>60.1</td>
<td>56.4</td>
<td>52.8</td>
<td>52.5</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>28.9</td>
<td></td>
<td>28.4</td>
<td>50.9</td>
</tr>
<tr>
<td>D</td>
<td>36.2</td>
<td>28</td>
<td>24.8</td>
<td>45.4</td>
<td>62.4</td>
</tr>
<tr>
<td>E</td>
<td></td>
<td>57.5</td>
<td></td>
<td>52.5</td>
<td>72.5</td>
</tr>
<tr>
<td>Chengdu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>55.9</td>
<td>55.9</td>
<td>55.1</td>
<td>60.1</td>
<td>38.1</td>
</tr>
<tr>
<td>B</td>
<td>70.2</td>
<td>62.4</td>
<td>62.8</td>
<td>52.3</td>
<td>35.3</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>26.6</td>
<td></td>
<td>28.4</td>
<td>41.3</td>
</tr>
<tr>
<td>D</td>
<td>22.5</td>
<td>27.1</td>
<td>26.1</td>
<td>31.7</td>
<td>48.2</td>
</tr>
<tr>
<td>E</td>
<td>22.9</td>
<td>25.2</td>
<td>24.8</td>
<td>48.2</td>
<td>48.2</td>
</tr>
<tr>
<td>Shanghai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>39.5</td>
<td>40.9</td>
<td>47.7</td>
<td>45.5</td>
<td>24.1</td>
</tr>
<tr>
<td>B</td>
<td>52.3</td>
<td>41.4</td>
<td>44.5</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>C</td>
<td>21.8</td>
<td>24.1</td>
<td></td>
<td>46.4</td>
<td>49.5</td>
</tr>
<tr>
<td>D</td>
<td>32.7</td>
<td>28.2</td>
<td></td>
<td>49.5</td>
<td>49.5</td>
</tr>
<tr>
<td>E</td>
<td>28.2</td>
<td>28.2</td>
<td>29.5</td>
<td>46.8</td>
<td>46.8</td>
</tr>
</tbody>
</table>

\(^1\)TL sums responses “much too little” and “too little”; TM sums responses “too much” and “much too much.”

---

**Table 3.** Statistics for partial least squares regression on dummy variables\(^1\)

<table>
<thead>
<tr>
<th>Sample</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beijing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>3.75</td>
<td>3.82</td>
<td>4.51</td>
<td>3.97</td>
<td>4.32</td>
</tr>
<tr>
<td>Observed mean</td>
<td>3.15</td>
<td>3.94</td>
<td>4.32</td>
<td>3.93</td>
<td>4.32</td>
</tr>
<tr>
<td>Predicted mean</td>
<td>0.60</td>
<td>0.78</td>
<td>0.19</td>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td>Overall mean drop</td>
<td>3.82</td>
<td>4.31</td>
<td>3.94</td>
<td>3.95</td>
<td>4.31</td>
</tr>
<tr>
<td>Chengdu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.90</td>
<td>2.75</td>
<td>3.79</td>
<td>3.74</td>
<td>3.90</td>
</tr>
<tr>
<td>Observed mean</td>
<td>2.90</td>
<td>3.05</td>
<td>3.79</td>
<td>3.74</td>
<td>4.20</td>
</tr>
<tr>
<td>Predicted mean</td>
<td>0.92</td>
<td>1.26</td>
<td>0.15</td>
<td>0.21</td>
<td>0.11</td>
</tr>
<tr>
<td>Overall mean drop</td>
<td>3.86</td>
<td>3.76</td>
<td>4.22</td>
<td>4.35</td>
<td>3.84</td>
</tr>
<tr>
<td>Shanghai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>3.28</td>
<td>3.28</td>
<td>3.92</td>
<td>3.86</td>
<td>3.71</td>
</tr>
<tr>
<td>Observed mean</td>
<td>3.28</td>
<td>3.28</td>
<td>4.07</td>
<td>4.26</td>
<td>3.71</td>
</tr>
<tr>
<td>Predicted mean</td>
<td>0.56</td>
<td>0.48</td>
<td>0.05</td>
<td>0.09</td>
<td>0.13</td>
</tr>
</tbody>
</table>

\(^1\)Observed mean: mean overall liking score of the consumers; mean drop: the difference between the mean overall liking score for the just-about-right group and mean overall liking score of the “too much” or “too little” categories.
results for Beijing. For all attributes except that of sample E, the dummy variables were significant, meaning that they were not at the just-about-right level. According to their increasing effect on liking, the most important attributes that need improvement for sample A were sweetness (−), fresh milk flavor (−), sourness (+), and thickness (−). The most important attributes for sample B were sweetness (−), sourness (+), fresh milk flavor (−), and thickness (−). The most important attributes for sample C were smoothness (+), sourness (−), fresh milk flavor (−), thickness (−), and thickness (−). The most important attributes for sample D were smoothness (+), fresh milk flavor (−), sourness (−), thickness (−), and sourness (−). For sample E, all mean drops of attributes were zero, meaning that they were on the just-about-right level.

Figure 7 indicates the PLS results for Chengdu. For all attributes except the sweetness of sample E, dummy variables were significant. The most important attributes needing improvement for sample A were fresh milk flavor (−), thickness (−), and sweetness (−). The most important attributes for sample B were fresh milk flavor (−), sourness (+), sweetness (−), thickness (−), and smoothness (−). The most important attributes for sample C were smoothness (+), thickness (−), fresh milk flavor (−), sourness (−), and sourness (−). The most important attributes for sample D were smoothness (+), thickess (−), sweetness (−), fresh milk flavor (−), and sourness (−). For sample E, all mean drops of attributes were zero, meaning that they were on the just-about-right level.

Figure 6. Partial least squares regression on dummy variable results of Beijing. Sample E is not shown because all mean drops of attributes were zero, meaning that there were on the just-about-right level. Mean drop: the difference between the mean overall liking score for just-about-right group and mean overall liking score of the “too much” or “too little” categories.
were smoothness (+), sourness (+), thickness (−), and fresh milk flavor (−). The most important attributes for sample E were smoothness (+), fresh milk flavor (−), and thickness (−).

**Comparison of Consumer Preference Between Cities**

Consumer acceptance of samples was evaluated by JAR scale, and significant differences in JAR levels
between cities were verified by $\chi^2$ analysis, the results of which are shown in Table 4. Values in bold indicate significant differences ($P \leq 0.05$) between cities for certain attributes of samples A to E. These results indicate that consumers from different cities have specific requirements for sourness, sweetness, and smoothness.

Potential improvements in liking (mean drops) of attributes were suggested by PA and PLS. Each attribute

![Figure 8](image-url)

**Figure 8.** Partial least squares regression on dummy variable results of Shanghai. Mean drop: the difference between the mean overall liking score for just-about-right group and mean overall liking score of the “too much” or “too little” categories.
was related to 2 mean drops (too little and too much), and the selection scheme of the significant attributes refers to Worch et al. (2010). The potential gain in liking was similar for attributes from PA and PLS on dummy variables, although the value of PLS on dummy variables was much smaller than that of PA. Most of the correlation coefficients between the estimated potential gain in liking for each attribute from both methods were high (>0.6).

These results indicate that consumers from 3 cities in China had different preferences for flavored liquid milk. Designing flavored liquid milk with these diverse location-specific requirements of 3 cities will result in higher overall liking scores for consumer-oriented products. China is a large country and citizens in different areas have particular preferences for food and taste characteristics. The regional differences of food culture are related to the geographical environment. According to the results of PA and PLS with dummy variables, consumers in Beijing would accept flavored milk with appropriate sourness and sweetness, and they preferred flavored milk where intensity of fresh milk flavor and thickness are high. Consumers of Chengdu preferred flavored milk with medium intensity of sourness and sweetness. Shanghai consumers preferred flavored milk with a high intensity of sweetness and did not have a specific requirement for sourness. Meanwhile, the high intensity of smoothness was not preferred by consumers in any of the 3 cities.

In consumer research and sensory sciences, the relationship between JAR variable and hedonic score is used to optimize or further develop a product. Just-about-right scales have been used to optimize salsas (Popper and Gibes, 2004), yogurt (Narayanan et al., 2014; Morell et al., 2015), raisin jams (Rababah et al., 2012), probiotic Petit Suisse cheese (Esmerino et al., 2013), mixed juices from Amazon fruits (Freitas and Mattietto, 2013), and juice blends (Lawless et al., 2013), among others. In addition to PA and PLS regression with dummy variables, several methods have been used to evaluate JAR data; for example, multivariate adaptive regression (Xiong and Meullenet, 2004), canonical variate analysis (Popper and Gibes, 2004), and Thurstonian ideal point modeling (Goerlitz and Delwiche, 2004). The main aim of analyzing JAR data is to identify the strengths and weaknesses of a given product and to determine which attributes should be increased or decreased in future product formulations.

Penalty analysis is probably the most commonly used method to analyze JAR data. However, for product developers, it would be very useful to identify in which direction a given attribute should be modified to affect the hedonic score the most (Gere et al., 2015). From the results reported in this study and in the literature (Xiong and Meullenet, 2006), PLS regression with dummy variables used a similar graphical presentation of relationships between JAR scale and hedonic liking variables. Unlike the PA method, PLS with dummy variables is a regression method that can reveal correlations among product characteristics and be used to predict consumer overall liking from JAR data; therefore, PLS regression can better estimate the mean drop of overall liking due to an attribute not being at the just-about-right level.

### CONCLUSIONS

Preferences for flavored liquid milk were diverse among consumers from 3 cities in China. For consumers in Beijing, appropriate levels of sourness and sweetness were important and consumers preferred flavored liquid milk in which the intensity of fresh milk flavor and thickness were high. Consumers in Chengdu preferred flavored liquid milk with medium intensity of sourness and sweetness and high intensity of fresh milk flavor and thickness. Consumers in Shanghai preferred samples with high intensity of sweetness and thickness, but did not have specific requirements for sourness. High intensity of smoothness was not particularly preferred for consumers of all cities studied. Overall, the results of penalty analysis and partial least squares regression were consistent. Both methods are product-specific, and improvement in the attributes of a product are denoted by the potential gain in liking (mean drops). The significant attributes selected by the 2 methods were similar and their mean drop scores were well cor-

---

**Table 4. Chi-squared analysis (P-values in parentheses) for significant differences (bolded) in just-about-right levels between cities**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Sample A</th>
<th>Sample B</th>
<th>Sample C</th>
<th>Sample D</th>
<th>Sample E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soursness</td>
<td>15.28 (0.004)</td>
<td>12.54 (0.014)</td>
<td>1.75 (0.782)</td>
<td>67.84 (0.000)</td>
<td>8.96 (0.062)</td>
</tr>
<tr>
<td>Sweetness</td>
<td>12.52 (0.014)</td>
<td>14.47 (0.006)</td>
<td>2.47 (0.650)</td>
<td>1.36 (0.852)</td>
<td>6.37 (0.173)</td>
</tr>
<tr>
<td>Fresh milk flavor</td>
<td>4.11 (0.391)</td>
<td>7.74 (0.102)</td>
<td>2.66 (0.617)</td>
<td>1.40 (0.845)</td>
<td>4.77 (0.311)</td>
</tr>
<tr>
<td>Thickness</td>
<td>6.18 (0.186)</td>
<td>5.78 (0.217)</td>
<td>8.14 (0.087)</td>
<td>2.23 (0.694)</td>
<td>5.10 (0.277)</td>
</tr>
<tr>
<td>Smoothness</td>
<td>9.59 (0.048)</td>
<td>7.24 (0.124)</td>
<td>3.56 (0.469)</td>
<td>10.22 (0.037)</td>
<td>5.89 (0.209)</td>
</tr>
</tbody>
</table>

1The difference analysis was conducted between cities. For example, the P-value of sourness for sample A is 0.004, indicating that significant differences are present between Beijing, Shanghai, and Chengdu for sourness of sample A.
related. Partial least squares on dummy variables is a regression-based method, which establishes a relationship between product characteristics and overall liking, whereas penalty analysis ignores the correlation. The design of flavored liquid milk should consider the diverse and location-specific requirements of consumers in China.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (No. 31201358), the Fundamental Research Fund for the Central Universities of China (06110607), and Science Foundation of Administration of General Administration of Quality Supervision, Inspection and Quarantine of the People’s Republic of China (2014QK220).

REFERENCES


