Food Quality and Preference 45 (2015) 1–10

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Quantitative analysis of product categorization in soft drinks using bottle silhouettes

Ana M. Arboleda [a,*], Carlos Arce-Lopera [b,1]

[a] Department of Marketing and International Business, Universidad Icesi, Cl 18 # 122-123, Cali, Colombia
[b] Department of Engineering, Universidad Icesi, Cl 18 # 122-123, Cali, Colombia

ARTICLE INFO

Article history:
Received 22 December 2014
Received in revised form 10 April 2015
Accepted 13 April 2015
Available online 20 April 2015

Keywords:
Beverage packaging
Categorization
Computer vision
Soft drinks

ABSTRACT

In our daily life, we use our senses to acquire information about the objects that surround us. However, the information processing that allows for the recognition and consecutive classification of those objects into categories remains unclear. Our purpose is to analyze the categorization mechanism taking into account: (a) package visual metrics and (b) consumer perceptions of this basic visual information. First of all, we quantitatively analyzed the physical characteristics of 52 bottle silhouettes of seven soft drink categories: sports drinks, water, flavored water, sodas, fruit juices, malt drinks and tea. We found that measures of the shape of the bottles can model the membership to a product category. Our first experiment tested how accurately consumers could recognize product category from real bottle silhouettes. We found that the visual characteristics that differentiate product category silhouettes are lid width and bottle shape (body kurtosis). Our second experiment tested the capacity of consumers to recognize artificially created bottle silhouettes. When basic information, such as the product shape is modified, consumers are not always capable of recognizing its corresponding category. We concluded that the physical attributes of bottles are related to the categorization process of the bottle content made by consumers. These findings may provide guidelines for new bottle designs that capitalize on existing categorization rules based on consumer perception.

1. Introduction

Product packaging is believed to enhance consumer awareness and brand recognition because not only is it informative about the product characteristics and attributes, but also has an esthetic meaning to the consumer (Bloch, 1995; Creusen & Schoormans, 2005). A package can be considered the most important communication tool of a product because it is the first and most accessible information that is available to consumers. It is most often the only decision-making tool available at the point of purchase and is the part of the product with which buyers are actively involved during the decision-making process and consumption (Ampuero, 2006).

The accuracy of the product categorization depends upon how the package is able to communicate its content and purpose; that is, the product category it represents. Thus, our purpose is to analyze the accuracy of the categorization mechanism providing information regarding both package design and consumer perceptions. There is extensive literature available regarding the cognitive categorization process in which consumers engage in choosing a product and making a final purchase decision (Bloch, 1995; Cohen & Basu, 1987; John & Sujan, 1990). Nevertheless, consumer behavior studies on categorization have evaluated this cognitive process based on consumer perceptions. We complement the consumer's perspective with the actual physical characteristics of package. In other words, before interpreting the consumer's categorization process, we confirm whether there is in fact a consistent product category according to visual package characteristics.

A package is defined as a product container that has a functional and symbolic purpose. The functions of packaging are to hold, protect, preserve, identify, and facilitate handling and sale of a product (Ampuero, 2006). Packaging has a symbolic purpose because, thanks to its shape and esthetic components, evokes particular product characteristics to the consumer. It plays an essential role in communicating product values as it is consciously designed to blend with a brand (Orth & Malkewitz, 2008). Finding the ideal packaging for a product is a challenge, but both designers and marketing managers should engage into this search for the most suitable kind of packaging for a specific product, considering all product performance requirements, regulations, and

* Corresponding author. Tel.: +57 2 5552334x8765.
E-mail addresses: amarboleda@icesi.edu.co (A.M. Arboleda), caarce@icesi.edu.co (C. Arce-Lopera).
[1] Tel.: +57 2 5552334x8529.

http://dx.doi.org/10.1016/j.foodqual.2015.04.006
0950-3293/

manufacturing restrictions (Bloch, 1995). It is difficult for product designers to determine which shape a product should have in terms of consumer perceptions to accomplish its informative and esthetic goals.

In particular, the consistency and accuracy of product informative characteristics should motivate product recognition and recall on the part of consumers. Based on product packaging, consumers should be able to determine the category to which a product belongs, its benefits, and form of use. At the point of purchase, consumers choose and decide to purchase a product based on their own experience, but if there is no previous experience and it is the first time a consumer makes a purchase decision, the majority (if not all) of the information is gathered from the packaging and labeling. Both experience and packaging characteristics have an influence on consumer perceptions of a product.

Nevertheless, to optimize the shopping experience, consumers rapidly review basic product characteristics and recognize its overall use and benefits. This cognitive process known as categorization (Bloch, 1995; Cohen & Basu, 1987) reduces the amount of information a consumer has to learn and understand during a shopping experience to be able to recognize a product as a member of a category. When a decision is made, an individual recalls basic cues that provide sufficient information to recognize a product. This paper evaluates packaging silhouette as a minimum packaging characteristic that furnishes information about a product category. We use this specific packaging representation to analyze the consumer's categorization process.

The literature on consumer categorization understands categorization as a cognitive process (Bloch, 1995; Cohen & Basu, 1987; John & Sujan, 1990). Also, the cognitive process depends upon the consumer's experience and upon the products available in the market. Therefore, this study evaluates consumer's categorization based on the physical characteristics of bottle silhouettes of soft drinks commonly available in the market. Consistently, our purpose is to analyze the categorization mechanism considering two levels at which this process takes place, namely, at the product and consumer level. We assume that accurate categorization firstly depends upon the consistency of products within a category and secondly upon an individual's capacity to recognize a product based on basic physical product attributes. Our literature review provides an understanding of the consumer's categorization mechanism as a cognitive process that facilitates the consumer's decision-making process.

Our objective requires two stages of analysis. First, we quantitatively analyzed the physical characteristics of 52 real bottles silhouettes of seven soft drink categories: sports drinks, water, flavored water, sodas, fruit juices, malt drinks and tea. The results show the consistency of the physical patterns of soft drink bottles in each category. Our second analysis is at the consumer level. We ran an experimental study that evaluated the consumer's ability to determine whether a product belongs to a particular soft drink category by using visual information about its silhouette (Gofman, Moskowitz, & Mets, 2010). Following, we tested the consumers' ability to recognize artificially created silhouettes. We manipulated the real silhouettes in two ways; firstly, by extracting an average silhouette for each category, and secondly, by creating a shrunken and expanded version of each silhouette. In this case, we also tested consumers' accuracy or ability to recognize each manipulated silhouette within the right category.

2. Categorization

In their everyday life, individuals use their vision to acquire information about the objects that surround them, like the freshness of vegetables (Arce-Lopera, Masuda, Kimura, Wada, & Okajima, 2015) or the age perception of the human skin (Arce-Lopera, Igarashi, Nakao, & Okajima, 2013); also, this includes learning about products and brands. However, not every object or product is noticed and memorized. The best individuals can do is to identify simple characteristics which are common to a product group, and to remember the group as a whole or as a category (Cohen & Basu, 1987). Categorization is a cognitive mechanism that enables consumers to understand a product by placing it within an existing category and to simplify their attention and decision-making processes (Bloch, 1995; Kahneman, 2011).

A product category represents a concept that groups a number of products that share similar attributes and characteristics. When there is not a sufficient number of attributes to define a category, boundaries tend to be fuzzy, and it is not always clear whether or not a product belongs to a particular category (Fiske & Taylor, 2008). The categorization process involves a comparison between different products to enable differentiation from others that belong to different categories. Subsequently, these differences are learned as functional and esthetic characteristics of a group. This process is important in a consumer environment because it allows individuals to find and choose a product easily by recognizing its basic attributes and benefits. Individuals use basic information like the shape or silhouette of a product to make an evaluation and recognize its attributes (Becker, van Rompay, Schifferstein, & Galetzka, 2011; Parise & Spence, 2012). Thus, a packaging shape that is easily assigned to a category is more efficient, compared to one that does not belong to a particular category, because consumers need less time and lesser details to understand its use and benefits. Hence, categorization is an important mechanism in the consumer's decision-making process.

The cognitive categorization process is explained through two alternative models: feature-based and exemplar-based models (Cohen & Basu, 1987). The feature-based model is based on a prototype perspective. The prototype perspective states that individuals reach a category prototype by finding an overlap between products of a category. The prototype is the central tendency or average of a product category (Fiske & Taylor, 2008). Details of every product are lost, and individuals are left with a schema-like representation that evokes a "typical" category product. When consumers find a new product, they learn about its characteristics and match these with a certain category.

The second model is the exemplar-based categorization perspective, which emphasizes the importance of some brands as individuals learn about a product category. Therefore, when individuals think of a category, a specific category member comes to mind. As the variation of product characteristics within a category increase, individuals tend to take the prototype approach. However, for more complex concepts the exemplar basis is a preferred path. In this case, prototypical features may be difficult to remember, whereas those of an exemplar are easily brought to mind. In practice, both perspectives depend upon contextual characteristics that influence choice and consumption due to an individual's exposure to particular brands through marketing and advertising (Cohen & Basu, 1987).

2.1. Study 1: market categories

At the point of purchase, products are organized into product categories, namely, sets of products that have similar functions and benefits. This preliminary study evaluates the extent to which product categories have a consistent shape, which is observed through product silhouettes. Consistency means that a given physical attribute has a similar measure across different kinds of bottles (i.e., relatively low variance). We choose the physical attributes of a bottle silhouette as the following: Centroid X coordinate (CX), Centroid Y coordinate (Emery, Kramer, & Tian, 2001), Body

Height (BH), Body Width (BW), Silhouette Area (SA), Lid Height (LH), and Lid Width (Westerman et al., 2013). This physical information is measured using bottles available in the market that conceptually form product categories. Thus, the following hypothesis is formulated:

H1. Product silhouette characteristics within a market product category have at least one physical attribute with low variance.

2.1.1. Stimulus acquisition and measurement

This preliminary section, which uses a computer vision approach, provides a quantitative analysis of the visual properties of soft drink bottle silhouettes to determine which physical measures are correlated with the real categorization of a product. This analysis is conducted for the following seven soft drink product categories: sodas, fruit juices, tea, water, sports drinks, flavored water, and malt drinks. Table 1 describes the seven soft drink categories.

Stimulus bottles were those available in the local market; likewise, the seven categories that represent different kinds of bottles were chosen from those existing in the local market for bottled soft drinks. Product classification within a category was determined by the actual content of the bottle, which was consistent with the classification of the product in the market. For example, if a bottle contained water, this was classified within the water category. If a bottle contained carbonated water, this was also classified within the soda category. All other contents are clearly classified within one of the seven soft drink categories.

2.1.2. Apparatus

High-resolution digital pictures of the front side of personal-sized bottles (less than 600 ml) were taken in a light-controlled environment (see Fig. 1). We used a Canon EOS7D camera, Canon Inc., Japan. We photographed 52 different bottles, corresponding to 70 different soft drink products; same bottle shape for various contents.

2.1.3. Stimuli

Using image processing techniques, the silhouettes of all bottle images were extracted, storing only the information of the bounding box circumscribed to the silhouette. Subsequently, seven overall physical properties of all bottle silhouettes were measured, namely, the Centroid X coordinate (CX), Centroid Y coordinate (Emery, et al.), Body Height (BH), Body Width (BW), Silhouette Area (SA), Lid Height (LH) and Lid Width (Westerman et al., 2013). These measurements were calculated using the Image Processing Toolbox of Matlab. Although other studies have also used product silhouettes or basic shapes as stimuli to evaluate consumer's associations (Becker et al., 2011; Parise & Spence, 2012), this study delves into further details of product shapes.

Fig. 2 shows schematics of the seven overall physical properties measured in pixel units. The centroid specifies the center of mass of the region. CX represents the horizontal coordinate of the center of mass. Similarly, CY is the vertical coordinate of the center of mass.

BH is the length in pixels between the bottom of the bottle and the start of the lid. BW is the length in pixels of the width of the bounding box. SA is the sum of the pixels in gray, and LH and LW are the vertical and horizontal lengths of the lid, respectively.

Additionally, we measured four statistical features of bottle silhouettes: Body Mean (BM), Body Variance (BV), Body Skewness (Westerman et al., 2013), and Body Kurtosis (BK). These measures were computed to add additional features that statistically encode bottle shapes. To obtain these statistical measures, the bottle body silhouettes were transformed into a distribution by rotating the body 90° counterclockwise and cutting the silhouette by half (see Fig. 3).

Subsequently, the first four central moments of the resulting distribution were calculated. The first central moment, the mean value, is a measure of central tendency. For the bottle distribution, the bottle mean is a measure of the overall bottle width which can be calculated using Eq. (1), where BM is the Bottle Mean, d the orthogonal distance from the center to the edge of the bottle, and n the body height in pixels:

Table 1
Description of each soft drink category.

Category name	Category description
Soda	Carbonated beverages, including carbonated water
Fruit juices	Beverages that contain fruits or liquids that are fruit flavored
Tea	Beverages containing tea
Water	Bottles that contain distilled, mineral, or spring water
Sports drinks	Beverages containing electrolytes to help the hydration process
Flavored-water	Products that add some flavor to water
Malt drinks	Brownish non-alcoholic beverages, brewed from barley, hops, and water

Fig. 1. Front view of the acquisition settings.

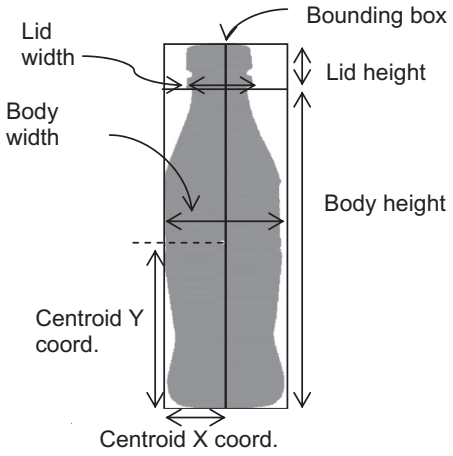

Fig. 2. Overall physical properties measured.

Fig. 3. Bottle silhouette distribution.

$$\mathrm{BM} = \frac{\sum_{i=1}^{n} d}{n} \tag{1}$$

The second central moment or variance is a measure of variability or dispersion of data above the mean. For the bottle distribution, the variance is an overall measure of variability with respect to the mean which can be formulated as Eq. (2), where BV is the body variance, d the orthogonal distance from the center to the edge of the bottle, BM the bottle mean, and n the body height in pixels:

$$\mathrm{BV} = \frac{\sum_{i=1}^{n} (d - \mathrm{BM})}{n} \tag{2}$$

The third central moment, skewness, is a measure of the level of asymmetry with respect to the mean. If the distribution is symmetric, then the skewness is zero. If a distribution has one tail that is longer than the other, it is skewed. If the longer tail is on the left side, it is then negatively skewed. If the longer tail is on the right side, it is positively skewed. For the bottle distribution, skewness can be formulated as Eq. (3), where BS is the body skewness, BV is the body variance, d the orthogonal distance from the center to the edge of the bottle, BM the bottle mean, and n the body height in pixels:

$$\mathrm{BS} = \frac{\sum_{i=1}^{n} (d - \mathrm{BM})^3}{n \cdot \sqrt{\mathrm{BV}^2}^3} \tag{3}$$

The fourth central moment or kurtosis is useful in testing univariate and multivariate normality; as a measure of departure from normality; in issues of robustness, outliers, and bimodality (DeCarlo, 1997). Tailedness and peakedness are both components of kurtosis basically because kurtosis represents a movement of mass that does not affect variance. Therefore, kurtosis is a statistical measure of the shape of a distribution. Eq. (4) shows the definition of kurtosis we used, where BK is the bottle kurtosis, BV is the body variance, d the orthogonal distance from the center to the edge of the bottle, BM the body mean, and n the body height in pixels:

$$\mathrm{BK} = \frac{\sum_{i=1}^{n} (d - \mathrm{BM})^4}{n \cdot \sqrt{\mathrm{BV}^2}^4} - 3 \tag{4}$$

Therefore, for each bottle silhouette, 11 physical measures were taken, namely, seven overall features and four statistical measurements.

2.1.4. Results

A linear discriminant analysis was conducted to determine which physical measures are relevant for the process of categorizing the seven product categories. The stepwise method for the linear discriminant analysis determined that four physical parameters (i.e. BH, BM, LH, and LW) suffice to create a categorization model of soft drinks for our local market. Therefore, three discriminant functions were developed. Table 2 shows the standardized coefficients of these discriminant functions.

Table 2
Standardized canonical discriminant function coefficients for market categories.

Parameter	Function		
	1	2	3
BH	−0.89	−0.53	2.52
BM	1.44	1.09	−1.90
LH	−0.44	0.76	−0.37
LW	0.79	−0.67	0.14
Exp. var.%	76.2	14.6	8.7

Table 3
Structure matrix for market categories.

Variables	Function		
	1	2	3
BK	0.625[a]	−0.030	−0.400
BS	−0.580[a]	−0.150	0.500
BM	0.600	0.683[a]	0.390
SA	0.320	0.599[a]	0.490
BW	0.080	0.503[a]	0.440
CY	−0.070	−0.462[a]	−0.340
BV	−0.240	0.330	0.735[a]
BH	0.410	0.610	0.678[a]
LH	−0.020	0.510	−0.070
LW	0.630	−0.280	0.020
CX	0.040	0.030	0.040

[a] Largest absolute correlation between each variable and any discriminant function.

The structure matrix (see Table 3) shows the correlation of each predictor variable with the discriminant function. BK and BS are the features that most strongly correlated with the first function. Similarly, BM, SA, BW, and CY are the features that most strongly correlated with the second function. Finally, a variability measure (BV) and the body height (BH) are most strongly correlated with the third discriminant function.

Given our data, the "null" model would be correct 30.7% of the time. The "null" model is when you choose the category with the largest size as the response for all stimuli. In our case, the largest category was the Soda category with 16 bottles. Therefore, the "null" model, i.e. selecting every stimuli as a member of the Soda category will be correct 16/52 = 30.7% of the time. Instead, our model has a correct classification rate of 82.7%. This results support hypothesis 1: product silhouette characteristics have consistent physical attributes that classify them within a given product category. Additionally, Table 4 shows that our model excels at identifying the malt drinks, sports drinks, and tea categories. Our worst scenario model results are for the flavored-water category with a 33.3% correct classification rate, which is still a better result in comparison to that of the "null" model.

2.1.5. Discussion

Before evaluating soft drink bottle categorization as a cognitive process, a preliminary analysis was performed to understand how soft drink bottles are classified in the real market. Our analysis was based on the physical characteristics of bottle silhouettes. The results of a discriminant analysis confirmed that simple physical measures of bottle shapes model the membership to a product category (a model classification rate of 82.7%). These results are also supportive of the idea that marketing managers, designers, and bottle manufacturers in the soft drink industry have created product categories that possess consistent physical properties. Specifically, the bottle shape (coded as BK and BS in our model) has proven to be the most important discriminant attribute, which accounts for 76.2% of the variance in the quantitative classification of soft drinks silhouettes.

Table 4
Classification results for market categories.

Category	Predicted group membership (%)							Total
	Malt drinks	Water	Flavored water	Sports drinks	Soda	Tea	Fruit juices	
Malt drinks	100.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0
Water	0.0	75.0	0.0	25.0	0.0	0.0	0.0	100.0
Flavored water	0.0	33.3	33.3	0.0	33.3	0.0	0.0	100.0
Sports drinks	0.0	0.0	0.0	100.0	0.0	0.0	0.0	100.0
Sodas	6.3	0.0	6.3	0.0	81.3	0.0	6.3	100.0
Tea	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0
Fruit juices	0.0	0.0	9.1	0.0	9.1	0.0	81.8	100.0

In our model, BK is related to bottle shape in terms of modality and presence of outliers. For example, the tea product category had the highest BK average, which means that, on average, tea bottles had silhouettes with sharper edges. Similarly, the soda category had the lowest BK average, which means that soda bottles had softer or rounder edges. On the other hand, BS is related to the top–bottom asymmetry of the bottle. The tea category had the lowest BS average, while the fruit juice category had the highest BS average. This means that the top side of tea bottles is relatively larger than that of fruit juice bottles in comparison with their respective bottoms. In other words, the ratio between the top and the bottom of the bottle body in the tea category is greater than that of the fruit juice category. Fig. 4 below shows some examples of silhouettes of members in the above mentioned categories.

This study shows that bottle shape is a physical property that can be used to model how products are organized to define a category. In the case of our model, the categories with the best classification results were malt drinks, sports drinks, and tea (see Table 4). These categories had specific, small variances in shapes. For the soda, fruit juice, and water categories, our model had a high classification rate (>75%). The difference is that these categories have a higher variance in certain physical attributes. For example, the water category had the highest variance in BK. In the case of the fruit juice category, the cumulative variance (i.e. the sum of feature variances) was the highest. However, for the soda category, the variance in shape of its members was among the lowest. This result could mean that there is an exemplar in this category that guides the shape of other products.

Finally, the flavored-water category scored the lowest classification rate with a classification rate of only 33%. Our model incorrectly classified flavored-water bottles in the water (33%) or the soda (33%) categories. The lack of distinctive shapes in this category may show that flavored water is the youngest category in terms of packaging design and positioning. Our analysis also revealed that this category had the lowest variance in shape (for both BK and BS features) proving that consistency in the shape of members in a category is not sufficient for correct categorization. Although the results for the flavored-water category are highly consistent in terms of bottle shape, this category is still far from having unique physical attributes.

In short, the basic design attributes of the soft drink bottles in the sample were created following distinctive product categories, which can enable consumers to recognize each category. These results on the categorization of soft drink bottle silhouettes are consistent with the market categorization. Therefore, this categorization model for soft drink bottles provides evidence that is consistent with the idea that marketers use observable patterns when they design a product for a category (Cohen & Basu, 1987). This preliminary analysis confirmed the following: (a) soft drink silhouettes can be classified into independent product categories, and (b) flavored-water bottle shapes are not merging as a consistent category. But are these categorization rules followed by actual consumers? To address this question, we designed the following experiment.

Fig. 4. Examples of silhouette images of soft drink categories.

2.2. Study 2: bottle silhouette categorization

After observing that soft drink bottles available in the market can be grouped into categories based on the similarity of their silhouettes, this second study analyzes how consumers categorize the same bottles using the same stimulus: silhouettes. The purpose of this experiment is to evaluate the extent to which consumers are able to recognize a soft drink category based on its silhouette. Hence, the following hypotheses are formulated:

H2. Using basic packaging information, namely product silhouette, consumers are able to assign a product package to a given category.

H3. Physical attributes of product categories in the market are those which are also used by consumers to categorize a product.

We conducted a psychophysical experiment to assess the consumer categorization process for our soft drink bottle silhouettes. Product categories were consistent with the preliminary section of the study: soda, fruit juice, tea, water, sports drinks, flavored-water, and malt drinks.

2.2.1. Participants

Subjects were 12 females and 14 males, for a total of 26 individuals who participated in the experiment (average age: 25.4 years; SD: 6.7 years). All subjects had normal or corrected-to-normal visual acuity and normal color vision. None of the subjects were experts in the sale or classification of bottles based on sensory evaluation. Additionally, no specific instructions were given. Written informed consent was signed after complete explanation of the study according to the regulation of the institutional ethics committee of the Universidad Icesi.

Fig. 5. Examples of original and silhouette images for two categories.

2.2.2. Apparatus and stimuli

The stimuli were displayed on a 14-inch LCD monitor of a Dell notebook computer in a dark room to prevent any surrounding light from interfering with the stimuli. As visual stimuli, we used the original bottle shape silhouettes from the previous experiment: soda, fruit juice, tea, water, sports drink, flavored-water, and malt. Silhouettes were displayed as if participants viewed the real product from a distance of 35 cm. Fig. 5 shows an example of the stimuli we used.

2.2.3. Procedure and data analysis

After a dark adaptation period of 10 min, subjects observed the stimuli displayed on the computer screen using both eyes. While looking at the images, subjects were asked to assign the bottle silhouettes to one of the following categories: soda, fruit juice, tea, water, sports drink, flavored water, and malt drink. The subjects assigned categories via radio buttons and confirmed their selection with a button using the computer mouse. Their response was named "perceived category". There were 156 trials (52 images × 3 repetitions) for each subject. Additionally, all images were displayed in a random order but there was not any restriction on stimuli presentation time. The experimental software was developed by the authors. Prior to the experiment, subjects responded a short questionnaire with the following control variables: age, gender, frequency of purchase of soft drink bottles, and consumption. Subjects were considered "expert consumers" in a category if they answered that they consumed at least one product in a given category every week. Subjects were not debriefed after the study.

For data analysis, we ran a linear discriminant analysis to determine which physical measures were relevant for the categorization process of our subjects. For each subject, the perceived category of each bottle silhouette was calculated using a three-trial approach. When the subjects classified the silhouette differently in the three trials, the perceived category was labeled as unknown. We computed accuracy of response as the percentage of subjects who accurately classified a given silhouette within a correct product category.

2.2.4. Results

The experimental results demonstrate that subjects were able to predict the bottle silhouette category accurately in 61.54% of the stimuli presented (hypothesis 2). Moreover, throughout all categories, the accuracy of response did not differ by age or gender. However, significant differences were observed by comparing consumer's expertise (frequency of purchase). Subjects classified as soft drink experts provided significantly different responses in terms of accuracy in comparison to those who were not experts, such that, expertise enhances the accuracy of product silhouette categorization ($t_{(1368)} = 4.21$, $p < 0.001$).

Table 5
Standardized canonical discriminant function coefficients for perceived categorization.

Parameter	Function			
	1	2	3	4
CX	−0.14	−0.14	0.41	0.87
CY	0.28	0.82	0.94	−0.15
BM	0.51	−0.10	0.87	−0.50
LH	−0.34	−0.23	0.69	−0.32
LW	0.89	−0.09	−0.40	0.52
Exp. var.%	65.4	16.1	10.2	6.3

The stepwise method of the discriminant analysis modeled the perceived category using five parameters (see Table 5). In comparison with our preliminary analysis, the parameters BM, LH, and LW remained in the model. However, BH was replaced with two additional parameters: CX and CY (see Table 5). Table 5 shows the standardized coefficients for the discriminant functions.

Supporting hypothesis 3, there are physical attributes of the silhouettes of products available in the market, which are used by consumers to categorize such products. Table 6, the structure matrix, shows that the first discriminant function is strongly correlated with lid width and body kurtosis. CY, BW, SA, BH, BM, and BV are strongly correlated with the second discriminant function. The third discriminant function does not correlate with any of the variables. However, the fourth discrimination function is strongly correlated with CX.

Table 6
Structure matrix for perceived categorization.

Variable	Function			
	1	2	3	4
LW	0.813[a]	−0.180	−0.060	0.260
BK	0.282[a]	0.010	0.130	0.190
CY	0.009	0.950[a]	0.220	0.110
BW	−0.080	−0.896[a]	−0.190	−0.230
SA	0.150	−0.866[a]	0.010	−0.280
BH	0.320	−0.640[a]	0.290	−0.320
BM	0.410	−0.603[a]	0.370	−0.330
BV	−0.110	−0.255[a]	0.190	−0.200
CX	−0.240	−0.250	0.430	0.802[a]
LH	−0.030	−0.480	0.390	−0.050
BS	−0.260	−0.060	−0.260	0.000

[a] Largest absolute correlation between each variable and any discriminant function.

Table 7
Classification results for perceived categorization.

Category	Predicted group membership							Total
	Malt drink	Water	Sports drink	Soda	Tea	Fruit juice	Unknown	
Malt drink	83.3	0.0	0.0	16.7	0.0	0.0	0.0	100.0
Water	0.0	88.9	0.0	11.1	0.0	0.0	0.0	100.0
Sports drink	0.0	0.0	100.0	0.0	0.0	0.0	0.0	100.0
Soda	0.0	16.7	0.0	77.8	0.0	5.6	0.0	100.0
Tea	0.0	0.0	0.0	0.0	100.0	0.0	0.0	100.0
Fruit juice	0.0	0.0	0.0	0.0	0.0	100.0	0.0	100.0
Unknown	0.0	0.0	0.0	0.0	0.0	0.0	100.0	100.0

The classification results in Table 7 show that, on average, the model correctly classifies 88.5% of all cases. In particular, it excels at classifying the sports drink, tea, fruit juice, and unknown categories (i.e., when a particular silhouette is unknown to the consumer).

Finally, an unexpected result was related to the flavored-water category. Subjects did not consider flavored water as the perceived category of any of the silhouettes they observed, although this was an option.

2.2.5. Discussion

The results show that consumers use the bottle shape and lid width as distinctive category features, accounting for 65% of the variance. Consumers do though use other characteristics such as bottle characteristics related to the object area and mass. However, these features were relatively less important in terms of explaining consumers' categorization, accounting for 16% of the variance. The predicted group membership of our model excelled at categorizing the sports drink, tea, and fruit juice categories. Following, the water, malt, and soda categories were modeled with 89%, 83%, and 78% precision, respectively. Consumers' perceptions revealed that flavored water is not a consistent category. In fact, this category was absent in consumer responses. Although the name "flavored water" was listed under every stimulus as a possible category option, this name was not attributed to any one of the silhouettes. The poor results of the flavored water category confirm that these bottles are not well positioned as a category.

The process of categorizing a product within a group should be something that consumers do without much thought (Kahneman, 2011). But if categorization does not occur easily, it is probably because consumers have not recognized the product and, consequently, have not realized that there is a group of products that has distinctive characteristics from others, but similar characteristics among them. If packaging characteristics are not consistent in a category, consumers will spend more time in their decision-making process, which is undesirable for soft drinks, a product that consumers rapidly recognize, purchase, and drink. Cognitive ease means that things are going well, there is no change and no need to redirect attention or make additional efforts (Kahneman, 2011). For example, the less variance across packaging in one category, the less effort an individual makes to recognize a kind of packaging in such category.

Recognizing the packaging of a product occurs rapidly, automatically, and effortlessly because it is coherent. Individuals are capable of associating different packages to one category as long as their characteristics are consistent. Consistency does not mean zero variance, it means having relatively similar packaging. Our preliminary analysis shows that even though packaging within a category is physically different, its characteristics are relatively consistent to be statistically classified as a category. However, flavored water is a good example of how low variance across products does not guarantee accurate categorization on the part of consumers. Therefore, consumer perceptions may be more sensitive to the variance on category shapes, confirming the importance of packaging consistency to communicate to consumers characteristics about a product that belongs to a particular category (Cohen & Basu, 1987; Silayoi & Speece, 2007).

Consumers are able to recognize a product in a category by remembering a prototypical product and knowing that, even though there are some changes across brands, there are features that communicate that packaging belongs to a specific category (Loken & Ward, 1990). Our results provide evidence that consistency in bottle shape and lid width can explain an important portion of the consumers' categorization process.

2.3. Study 3: artificial bottle silhouette categorization

In study 2, a high-level cognitive decision task, the process of categorizing bottles was modeled with a linear regression model using five physical measures as parameters. This third study replicates the previous experiment modifying the original silhouettes. This experiment has the same apparatus and procedure as those used in the previous experiment (study 2). Also the same subjects participated.

To test further the hypothesis that shape is a crucial factor for categorization, we created artificial stimuli, which represented modified versions of the stimuli in study 2:

H4. Consumers perceive the average silhouette to represent the product category.

H5. By modifying the average silhouette, consumers are not able to assign a product packaging to a given category.

2.3.1. Stimuli

Using image-processing techniques, three different types of stimuli were created for each category. First, the average silhouette was computed using all bottle silhouettes in a given category. Then, we modified that average silhouette by shrinking it using a value of 1 SD, resulting in the shrunken images. Finally, in a similar manner, we added 1 SD to the average silhouette, resulting in the expanded images. Therefore, 21 artificial silhouettes were created: seven categories by three images per category (see examples on Fig. 6).

2.3.2. Data analysis

Similar to study 2, for each subject, the response about the perceived category of each bottle silhouette was calculated using the three-trial approach. When the subjects classified the silhouette differently for the 3 trials, the perceived category was labeled as

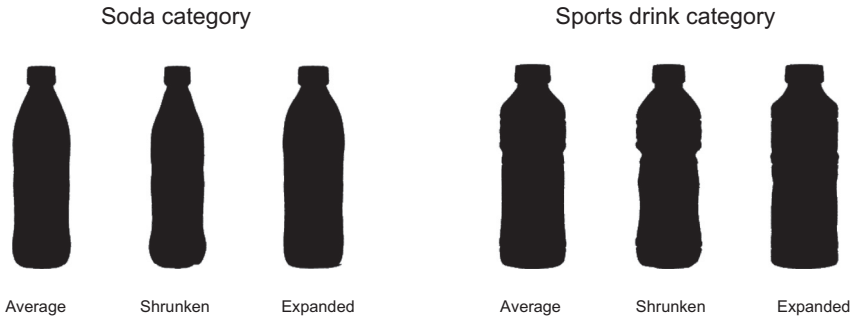

Fig. 6. Examples of artificial images for two categories.

Table 8
Highest percentage of classification with associated category.

Categories	Mean	Shrunken	Expanded
Water	Water (67.9)	Water (65.4)	Water (65.4)
Sports drink	Sports drink (67.9)	Sports drink (79.5)	Sports drink (61.5)
Soda	Soda (39.7)	Soda (76.9)	Water (50.0)
Tea	Tea (47.4)	Tea (46.2)	Sports (48.7)
Fruit juice	Water (39.7)	Fruit juice (23.1)	Tea (43.6)
Flavored	Soda (47.4)	Soda (47.4)	Soda (37.2)
Malt drink	Fruit juice (34.6)	Fruit juice (30.8)	Tea (25.6)

unknown. We computed the percentage of correct classification for each artificial image (i.e., accuracy of response).

2.3.3. Results

Table 8 shows the experimental results for each artificial image. The columns represent the three image transformations, and the rows represent the seven soft drink categories. Therefore, the column–row intersection shows the category that most subjects selected when they were asked to assign a category to a particular silhouette. The percentage of subjects who chose a category is shown in parenthesis. Hypothesis 4 is supported for four categories: water, sports drink, soda, and tea. In these four cases, the average silhouette accurately represents the product category.

Hypothesis 5 tested the extent to which consumers were confused by the modification of the silhouette. The hypothesis was not supported for water and sports drinks. When their shapes were artificially modified, consumers were still able to classify these silhouettes into their respective categories correctly (see Table 8). Hypothesis 5 was partially supported for soda, tea, and juice. The artificially shrunken silhouette was accurately classified in the case of soda, tea, and juice, but when the silhouette was expanded, these three products were mistaken for water, sports, and tea (respectively).

Finally, hypothesis 5 was totally supported in the case of artificial images created for the flavored water and malt drinks. Images for flavored water were consistently misclassified as members of the soda category. The subjects misclassified the artificial silhouettes of the malt drink category as members of the fruit juice or tea category.

2.3.4. Discussion

We tested the consistency of the categorization process using artificially modified bottle silhouettes. This technique evaluated if individuals were able to assign a product to its category by using an average silhouette (hypothesis 4). This is consistent with the idea that consumers create a prototypical image of a product category by comparing and overlapping the physical characteristics of products (Cohen & Basu, 1987).

However, when basic information, such as the shape of the product, is modified from reality, consumers are not always capable of recognizing its corresponding category and they are not as mistaken as we expected. This is particularly important for new products or traditional products that change their packaging. The best scenario for categories that are well known in the market is to design packaging that is close to the average shape so that consumers are able to recognize this product within its product category.

3. General discussion

This study compares consumers' categorization accuracy (cognitive process) with actual categorization of soft drinks given packaging physical characteristics (market categories). The comparison of actual physical characteristics with consumer perceptions is a unique method for analyzing the process of categorization, as it requires an understanding of consumer behavior through the methods of computer vision technology. On the one hand, consumer behavior literature explains individual preferences and purchase decisions after a cognitive process. On the other hand, computer vision evaluations use technology to represent an object in order to analyze individual perceptions, for example, identifying characteristics of an object or recognizing the identity of an object (Gunasekaran, 1996).

The use of computer vision provides a different approach to the consumer behavior literature in understanding the consumers' categorization process because the analysis of product attributes begins with the product itself. Human perception of such attributes comes later when individuals learn about a product. By using image-processing techniques, product characteristics are quantitatively analyzed, extracting visual information such as size and shape. This information is then related to consumer perception and decision-making about the product (Du & Sun, 2006). Other studies using computer vision technology also analyze color and texture.

Through a computer vision approach, we extracted quantitative information about the most basic image that represents a product, its silhouette. Because silhouettes represent products in a basic, general way, this study makes it easier to keep track of other products and contextual variables. Subjects do not have information related to the brand, color, nutrition facts, or any other information on the packaging or labeling, or information commonly observed at the point of purchase. Here, the bottle silhouettes are the only information individuals have to identify a product category.

Consequently, our findings suggest some guidelines for new bottle designs that capitalize on existing categorization rules based on consumer perception. At the point of purchase a product should be easily recognizable and able to communicate its identity in the most efficient way (Rettie, 2000). By understanding the physical characteristics of silhouettes in each category that provide more

consistent information to consumers, package designers and marketers would be able to determine which characteristic should they use to represent a product within a category accurately and which ones not to use causing misunderstanding of the product.

A quick and easy categorization process is important for products that are consumed frequently and do not represent a meaningful purchase. Soft drinks are a good example of this kind of product, where individuals want to minimize the effort they put into choosing and purchasing a beverage (Creusen & Schoormans, 2005). There is a trade-off between introducing a different characteristic to packaging that belongs to a certain category and tapping into the benefits of that category (Schoormans & Robben, 1997). A new stimulus does demand more attention from a consumer, but this same stimulus warns the consumer that the product is different from the existing product category. For this reason, it is important to determine which physical characteristics are more appropriate to enable an accurate consumer decision-making process.

4. Contributions

Through a set of studies, this paper presents some metrics for assessing the categorization of consumer products. These metrics enable for a direct correlation of the categorization process with physical attributes of the bottle silhouettes. Supporting the consumers' cognitive categorization process, we find that consumer perception of a product category is related with the actual physical measures of the products. This is important because although some similarity across products within a category is important for consumers to be able to recognize a product, some variance is also acceptable, and individuals should be able to understand those differences without misinterpreting product attributes. This is particularly relevant in the case of new or hybrid products that combine attributes of different categories (Rajagopal & Burnkrant, 2009).

These findings permit the construction of guidelines for new bottle designs that capitalize on existing categorization rules based on consumer perception. Finally, our study shows that the consumers' categorization process requires basic, simple information, just like a product silhouette. With this essential stimulus consumers are able to recognize a product category.

5. Limitations

We must notice that a packaging category that is recognized by the consumer through its silhouette is not necessarily the best kind of packaging. It only means that this is a category type of packaging either as an "average" of the category or as an exemplar. Packaging that represents the category is easy to remember and recognize, but it is not the best in terms of uniqueness. For innovative products that do not actually belong to a mass consumption type category, consumers might need to spend more time and cognitive resources to understand the use and benefits of a product. It is important to understand what product characteristics drive the consumer categorization process. Thus, a new packaging design should first consider what it needs to be easily classified as a product within a certain category, and second, what details can be shifted to signal brand differences. Our research does not include extreme silhouettes, thus we cannot predict what would happen in terms of consumers' categorization when a product has a significant departure from the category prototype. An example of an extreme silhouette is when roundness is accentuated, getting closer to the shape of a fruit (see for example bottle posted by (Hyden., 2012).

The scope of this paper is limited to silhouette as a source of information and a basic categorization element. At the point of purchase, consumers do have more information to recognize products within a category and to evaluate their attributes. In a natural setting like a supermarket, consumers have more information on the packaging: brand, label, display, and point of purchase advertisements and promotions. These visual stimuli are sensory information that facilitate consumers' associations and retrieval (Parise & Spence, 2012; Westerman et al., 2013). We limited the study to silhouette stimuli because we intended to show that by using basic, simple information consumers would be able to recognize a product category and differentiate it from others. Therefore, this set of studies explains the consumers' ability to categorize based on bottle silhouettes. On the down side, we did not evaluate the extent to which other characteristics like colors, names, or logos, facilitate product categorization by consumers. However, our results proved that silhouettes are important and may indicate that when the consumer walks down the supermarket isle, a quick glance at products only provides the brain with a blurred, silhouette-style image of the product. And maybe if a product is not recognized quickly, it is ignored.

Another external validity issue is the restriction to a set of categories defined by the market. The seven categories included in the study were those available in the local market for bottled soft drinks. For example, there is not an actual category for sparkling water; in our local market, there is only one brand of this type, and it is offered in the soda category. Although the seven selected categories are fairly common and general, in other contexts the set of categories may change.

6. Conclusion

This paper presents metrics for assessing the categorization of consumer products based on simple physical characteristics. For the case of soft drink bottles, the lid width and the body kurtosis are important visual cues for categorization. Finally, consumers are usually able to recognize a product category using basic information such as a product silhouette, even if the original silhouette is modified.

References

Ampuero, O. (2006). Consumer perceptions of product packaging. *The Journal of Consumer Marketing, 23*(2), 100–112.

Arce-Lopera, C., Igarashi, T., Nakao, K., & Okajima, K. (2013). Image statistics on the age perception of human skin. *Skin Research and Technology, 19*(1), e273–e278.

Arce-Lopera, C., Masuda, T., Kimura, A., Wada, Y., & Okajima, K. (2015). Model of vegetable freshness perception using luminance cues. *Food Quality and Preference, 40*, 279–286.

Becker, L., van Rompay, T. J., Schifferstein, H. N., & Galetzka, M. (2011). Tough package, strong taste: The influence of packaging design on taste impressions and product evaluations. *Food Quality and Preference, 22*(1), 17–23.

Bloch, P. H. H. (1995). Seeking the ideal form: Product design and consumer response. *The Journal of Marketing, 59*(3), 16–29.

Cohen, J. B., & Basu, K. (1987). Alternative models of categorization: Toward a contingent processing framework. *The Journal of Consumer Research, 13*(4), 455–472.

Creusen, M. E. H., & Schoormans, J. P. L. (2005). The different roles of product appearance in consumer choice. *Journal of Product Innovation Management, 22*(1), 63–81.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Method, 2*(3), 292–307.

Du, C.-J., & Sun, D.-W. (2006). Learning techniques used in computer vision for food quality evaluation: A review. *Journal of Food Engineering, 72*(1), 39–55.

Emery, C., Kramer, T., & Tian, R. (2001). Customers vs. products: Adopting an effective approach to business students. *Quality Assurance in Education, 9*(2), 110–115.

Fiske, S., & Taylor, S. (2008). *Social cognition, from brains to culture.* New York: McGraw-Hill Humanities.

Gofman, A., Moskowitz, H. R., & Mets, T. (2010). Accelerating structured consumer-driven package design. *The Journal of Consumer Marketing, 27*(2), 157–168.

Gunasekaran, S. (1996). Computer vision technology for food quality assurance. *Trends in Food Science & Technology, 7*(8), 245–256.

Hyden. (2012). Drinkable Review Retrieved February, 10, 2015, from <http://www.drinkablereview.com/2012/03/pom-wonderful-100-pomegranate-juice.html/>.
John, D. R., & Sujan, M. (1990). Age differences in product categorization. *Journal of Consumer Research, 16*(4), 452–460.
Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Straus and Giroux.
Loken, B., & Ward, J. (1990). Alternative approaches to understanding the determinants of typicality. *Journal of Consumer Research, 17*(2), 111–126.
Orth, U. R., & Malkewitz, K. (2008). Holistic package design and consumer brand impressions. *Journal of Marketing, 72*(3), 64–81.
Parise, C. V., & Spence, C. (2012). Assessing the associations between brand packaging and brand attributes using an indirect performance measure. *Food Quality and Preference, 24*(1), 17–23.
Rajagopal, P., & Burnkrant, R. E. (2009). Consumer evaluations of hybrid products. *Journal of Consumer Research, 36*(2), 232–241.
Rettie, R. (2000). The verbal and visual components of package design. *The Journal of Product & Brand Management, 9*(1), 56–70.
Schoormans, J. P. L., & Robben, H. S. J. (1997). The effect of new package design on product attention, categorization and evaluation. *Journal of Economic Psychology, 18*(2–3), 271–287.
Silayoi, P., & Speece, M. (2007). The importance of packaging attributes: A conjoint analysis approach. *European Journal of Marketing, 41*(11/12), 1495–1517.
Westerman, S., Sutherland, E., Gardner, P., Baig, N., Critchley, C., Hickey, C., et al. (2013). The design of consumer packaging: Effects of manipulations of shape, orientation, and alignment of graphical forms on consumers' assessments. *Food Quality and Preference, 27*(1), 8–17.