Food Quality and Preference 32 (2014) 83–92

Contents lists available at SciVerse ScienceDirect

Food Quality and Preference

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

External preference segmentation with additional information on consumers: A case study on apples

E. Vigneau [a,b,*], M. Charles [a,c], M. Chen [a,b]

[a] LUNAM University, ONIRIS, Sensometrics and Chemometrics Laboratory, Nantes, France
[b] INRA, Nantes, France
[c] LUNAM University, Groupe ESA, UPSP GRAPPE, Angers, France

ARTICLE INFO

Article history:
Received 7 September 2012
Received in revised form 3 April 2013
Accepted 14 May 2013
Available online 29 May 2013

Keywords:
Hedonic study
Segmentation
Consumer attributes
L-shaped data
Apples

ABSTRACT

We consider hedonic studies when, in addition to liking scores, external information is available on the products (i.e. sensory descriptors) as well as on the consumers (demographic, usage and attitude attributes). The classification around latent variables (CLV) methodology may be used for segmentation purposes in such situations. Two alternative strategies have been compared on the basis of a case study on 31 apple varieties according to the use *a priori* or *a posteriori* of the consumer attributes. The direct approach, L-CLV, which involves the three blocks of information (product hedonic scores, product sensory descriptors and consumer attributes) simultaneously, has demonstrated its ability to reveal a segmentation of consumers associated with a large number of sociological and behavioral parameters, in relation to the key sensory drivers. On the contrary, using a two-step procedure, with first an external preference segmentation by taking into account only the external information on products, no relevant information was gained with the subsequent use of the consumer attributes. For a better investigation of consumer preferences from a marketing research point of view, it appears that it is much more relevant to introduce both types of external information simultaneously and that L-CLV is suitable for this purpose.

1. Introduction

External preference mapping is a very popular methodology which aims to provide information about the main "drivers of liking" of consumers regarding the sensory (or physico-chemical) properties of the products of interest (Danzart, Sieffermann, & Delarue, 2004; Greenhoff & MacFie, 1994; Meullenet, Xiong, & Findlay, 2007; Naes, Brockhoff, & Tomic, 2010; Van Kleef, van Trijp, & Luning, 2006). It attempts to relate the sensory profile data to consumer liking scores using various standard statistical methods. The first step of the methodology is to create latent variables based on product sensory attributes. Usually, the two first principal components of the sensory data are considered but other strategies have also been proposed (Faber, Mojet, & Poelman, 2002; Plaehn, 2009; Verdun, Cariou, & Qannari, 2012). Thereafter, these sensory latent variables are used to model the individual consumer likings by means of linear models of varying complexity (vectorial, circular, elliptical or quadratic models). Alternatively, instead of modeling each individual separately, segments of consumers with relatively homogeneous acceptance patterns may be considered. This makes it possible to summarize the hedonic data by the average in each segment. The segment models are finally fitted separately on the sensory latent variables. However, in this process, segmentation and modeling are achieved separately.

It seems more relevant to merge consumers who have similar drivers of liking, rather than to identify segments of consumers having similar acceptance patterns and, afterwards, interpret these patterns in the light of the sensory attributes of the products. In order to define groups of consumers and, simultaneously, in each group, the prediction model of the liking scores as a function of the sensory attributes, a segmentation approach which takes account of external data on the products was proposed by Vigneau and Qannari (2002). In practice, this was achieved using the clustering around latent variables (CLV) approach (Vigneau & Qannari, 2003; Vigneau, Qannari, Sahmer, & Ladiray, 2006). This methodology of variables clustering offers the possibility of defining "directional" or "local" groups, of introducing co-variables measured on the same samples and/or additional information on the variables to be clustered themselves. In preference mapping, the L-CLV (an extension of the CLV for L-shaped data) may be applied in order to identify the drivers of liking in segments of consumers and also characterize these segments in terms of demographic, usage and attitude variables collected on the consumers (Vigneau, Endrizzi, & Qannari, 2011).

* Corresponding author at: LUNAM University, ONIRIS, Sensometrics and Chemometrics Laboratory, Nantes, France. Tel.: +33 2 51 78 54 40; fax: +33 2 51 78 54 38.
E-mail address: evelyne.vigneau@oniris-nantes.fr (E. Vigneau).

0950-3293/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodqual.2013.05.007

It should be noted that, using a rather different framework, latent class vector models (De Soete, & Winsberg, 1993, Courcoux, & Chavanne, 2001) may be adapted for the inclusion of covariates, such as the sensory attributes of the samples (Poulsen, Brockhoff, & Erichsen, 1997). However, the estimation of the parameters of these models requires the implementation of relatively complex EM-algorithms. With algorithms somewhat comparable to those of the CLV, another approach has also been developed in order to merge consumers who have similar drivers of liking. This is based on the fuzzy C-means (FCM) methodology and uses the residual distance between the linear combination of the sensory descriptors and the individual likings (Berget, Mevik, & Naes, 2008; Bolling Johansen, Hersleth, & Naes, 2010; Menichelli, Olsen, Meyer, & Naes, 2012; Wedel & Steenkamp, 1991). In addition to fuzziness, with the membership values of each consumer ranging between 0 and 1, this latter approach has the advantage of allowing an incomplete design in which each consumer has not tested all the samples. Nevertheless, additional information on the consumers cannot be directly introduced into the optimized criterion. Usually (Delgado, & Guinard, 2012; Helgesen, Solheim, & Naes, 1997; Naes, Kubberod, & Sivertsen, 2001; Sveinsdóttiret al., 2009; among others), the segments are related *a posteriori* to demographics or other consumer attributes by some type of linear regression analysis, PLS or PCR regression or factorial discriminant analysis. In the conjoint analysis context, Naes et al. (2010) have compared a simultaneous analysis, combining experimental factors and consumer attributes, with two-step approaches based on the PLS or PCR regression. However, the proposed simultaneous approach, using conventional linear models, requires unfolding the data matrices and selecting a limited number of consumer attributes (typically age group, gender or a categorical variable associated with some groups of consumers defined beforehand). Until now, only the L-PLS regression (Martens et al., 2005) or the L-CLV (Vigneau et al., 2011) have enabled the three blocks of data available (liking scores, product attributes and consumer attributes) to be analyzed in a single step, without unfolding and regardless of the amount of additional information collected on the consumers. The L-PLS regression results in graphical representations where the factorial components are defined according to the relationships between all types of variables whereas the L-CLV is oriented towards the segmentation of the panel.

We will focus on the advantages or disadvantages of performing the segmentation of a panel of consumers using the L-CLV method compared to a two-step approach. In fact, our objective is to understand better how slightly different data analysis strategies impact the interpretation of the data. The L-CLV represents a direct approach simultaneously involving the liking scores and the personality information on the consumers as well as the additional variables measured on the samples. The two-step approach consists of an external preference segmentation in which the consumer attributes are not part of the primary step of data analysis but are related afterwards to the results of the segmentation. The primary step is performed using the CLV algorithm by taking into account only the additional data on the products.

Both of these strategies will be compared on the basis of a recent experiment conducted on apples. A large set of apple varieties was described, in parallel, by a trained sensory panel and rated in terms of liking by consumers. In addition, consumers were asked to fill in a questionnaire.

2. The case study

A study on 31 batches of apple varieties produced in France (the Loire valley) was conducted at the beginning of 2011. The products were chosen in order to cover as much variability as possible in the texture, taste and aroma that can be found on the apple market. The product set was composed of well-known commercial apple varieties, new varieties and more rustic ones. These are listed in Table 1.

Table 1
List of the 31 apple varieties.

Code	Variety
ARI	Ariane
ARI2	Ariane2
BC	Belchard Chantecler
CHA	Chailleux
CM	Caméo
COX	Cox's Orange Pippin
CRI	Crimson Cripps
DJU	Delbard Jubilé®
DLC	Dalincot
DLS	Dalinsweet
DLT	Dalitron
FJ	Fuji
GD	Golden Delicious
GR	Goldrush
GR2	Goldrush2
GS	Granny Smith
GSA	Golden de Savoie
HC	Honey Crunch®
HC2	Honey Crunch®2
JAZ	Jazz™
JON	Jonagored
JUL	Juliet
PIL	Pink Lady®
PIN	Corail® Pinova
RA	Reinette d'Armorique
RBR	Reinette de Brive
RCLO	Reinette Clocharde
RG	Royal Gala
RGC	Reinette grise du Canada
SW	Schneywell
TT	Tentation®

A trained panel of 15 assessors was selected and trained according to ISO standards 8586 (ISO, 1993) and 11035 (ISO, 1995). They agreed on a list of 30 attributes including 19 descriptors for aroma. After a statistical analysis based on the citation frequencies and the redundancy among the attributes in each category (texture, flavor, aroma), 15 attributes were finally retained: three for the texture (crunchy, juicy, fondant), two for the flavor (sweet, acid), the overall odor intensity, the overall aroma intensity and eight descriptors for specific aromatic notes (A_Pineapple/Banana, A_Sweet/Rose, A_Woody/Earthy, A_Rustic, A_Lemon, A_Whiteflowers, A_Ripe fruit, A_Green). All attributes were evaluated on a 10 cm unstructured scale anchored from 0, not perceived, to 10, extremely intense. Products were presented in a monadic way according to a balanced design to avoid order and carry-over effects. Five eighths of each apple variety were served to each judge. Products were evaluated in duplicate. Profile measurements were carried out in sensory computerized booths according to NF ISO 8589 standards. Scores were collected with FIZZ (version 2.10; Biosystems, Courtenon, France). The sensory room was kept at 21 ± 1 °C, red lights were used and rinsing with mineral water between samples was mandatory.

During the same period of time, 224 regular apple consumers were recruited locally. The panel was balanced for gender, active/inactive people and for four age categories (18–25, 26–40, 41–55 and 56 years old and over). Products were presented monadically, in a random order, at room temperature. A blind warm-up sample was presented at the beginning of each session to avoid an effect of the first product (Wakeling, & MacFie, 1995). For the evaluation of each product, two eighths of peeled apple were served under white light. The test took place over four consecutive weeks. As far as possible, each consumer participated in one session per week.