Food Quality and Preference 66 (2018) 127–140

Contents lists available at ScienceDirect

Food Quality and Preference

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

Effects of adding extra samples to a product set when using descriptive analysis

Pauline Lestringant[a,*], Julien Delarue[b], Hildegarde Heymann[a]

[a] Departments of Food Science & Technology and Viticulture & Enology, University of California, Davis, USA
[b] AgroParisTech, UMR 1145 Ingéniérie Procédés Aliments, F-91300 Massy, France

ARTICLE INFO

Keywords:
Descriptive analysis
Product addition
Choice of products
Similarities
Differences

ABSTRACT

When performing descriptive analysis, it could be useful to add extra products to a product set when one only has few products or when comparing prototypes to specific reference products. Yet, adding products could modify the attributes generated and the context of evaluation, and result in differences in the description of the target set. This study investigates the effect of adding products to a target set, considering both the type and number of products added.

Each experiment was done on two product categories, ketchup and lemonade, to check the reproducibility of conclusions. Panels were trained on either four or eight products, to understand how that impacted the attributes generated. To evaluate the effect of the type of product added on ratings, two products were added to the four training products during evaluation, that were either similar or different from those. The effect of the number of products was investigated by adding either four or six products. Thus, six different panels were conducted. Statistical analyses were performed on the datasets reduced to the four target products.

Overall, the product structure was well maintained despite the different evaluation contexts and differences between product categories. However, separate analyses showed that product space complexity increased with the number of products. Moreover, the most salient descriptors on the first two dimensions varied depending on the evaluation context.

This study shows that it is possible to add products to a target set but the experimenter should keep in mind that it may modify the product descriptions.

1. Introduction

Descriptive analysis (DA) is one of the major methods used to describe the sensory characteristics of products. There are a number of variants of descriptive analysis, depending on the recruitment, training, data collection and statistical analysis methods used (Delarue, 2014; Lawless & Heymann, 2010; Murray, Delahunty, & Baxter, 2001). The following aspects of recruitment and training have been extensively discussed and studied in the literature, either on descriptive analysis or on rapid descriptive methods: number of panelists (Heymann, Machado, Torri, & Robinson, 2012), screening (Issanchou, Lesschaeve, & Köster, 1995; Zamora & Guirao, 2004), length of training (Wolters & Allchurch, 1994), attribute generation (Lawless & Heymann, 2010), use of standards (Rainey, 1986), training on the use of scales (Meilgaard, Civille, & Carr, 2006). In terms of data collection, methods vary on the types of scales used (Jeon, O'Mahony, & Kim, 2004), the number of replicates performed by each panelist, as well as whether a computer based or paper based system is used (Savidan & Morris, 2015). Depending on the goals of the study, an experimenter may choose different combinations of the above parameters. General good practices and common misuses of such methods can be found in the literature (Lawless & Heymann, 2010; Sidel, Stone, & Bloomquist, 1981). However, the literature rarely specifies an optimal number of products, or the strategies to pick a product set for generic descriptive analysis studies. In fact, many experimenters may not have the option to choose the number of products to include in their studies. Other researchers may not even consider that adding extra products may have an impact on the conclusions.

In contrast, the issue of the number of products to use for rapid methods is addressed in the literature, and determined mostly based on memory and sensory fatigue, due to either the task or the type of product used (Varela & Ares, 2014).

Having few products is a frequent case in industry or academia: when using a 2 × 2 experimental design, comparing a single prototype to a control, or when testing innovative products for which few competitors are on the market. Yet, when the study focuses on few products

* Corresponding author at: 2003 Robert Mondavi Institute – Sensory Building, University of California – Davis, One Shields Avenue, Davis, CA 95616, USA.
E-mail addresses: plestringant@ucdavis.edu (P. Lestringant), julien.delarue@agroparistech.fr (J. Delarue), hheymann@ucdavis.edu (H. Heymann).

https://doi.org/10.1016/j.foodqual.2018.01.017
Received 1 November 2017; Received in revised form 19 January 2018; Accepted 24 January 2018

(less than 6), issues can arise at several points. Panelists could memorize the products during training and rely more on memory than perception for their evaluation (Lestringant, Delarue, & Heymann, 2017b). In addition, the use of multivariate data analysis to create product maps with few products would not be adequate (Ares & Jaeger, 2014). A potential solution to this problem would be to artificially increase the sample size by adding extra products to the set. The experimenter would then go back to a more usual situation with more than 6 products (Lestringant, Delarue, & Heymann, 2017a). Adding products to a target set of products is also useful when comparing studies in time and building a cross-study database. In this case, the researcher will seek comparison points by choosing one or more products, stable in time, to add to each product set. This is for example the case of methods such as Polarized Sensory Positioning (PSP, Teillet, 2009), or Pivot Profile (Thuillier, 2007).

However, adding extra products modifies the context of evaluation and could impact the descriptions of the target products. Theoretically, some methods, such as Spectrum (Meilgaard et al., 2006), are not context-dependent as panelists are trained on physical standards matching different levels of intensity of the attribute scale. Practically, during evaluation, many context effects may occur. This could for example depend on the sequence of samples presented to a panelist. Presentation orders, such as William's Latin square designs, are one way to control for sequence effects such as assimilation or contrast (Lawless & Heymann, 2010). Other effects due to the overall spread of the product set cannot be mitigated through presentation orders. This is for example the case of range and frequency effects (Parducci, 1963), or centering biases (Poulton, 1989). According to these authors, panelists tend to center their ratings around a middle product and adjust their ratings so that the whole scale is used over the presented product set. Adding products, especially at the border of the sensory space would rescale all the ratings for other products. Moreover, if many products were added in a specific part of the scale, panelists would shift the center of the scale towards that end and would spread their ratings more in that area to avoid superimposing many products at one end (Parducci, 1974; Riskey, Parducci, & Beauchamp, 1979). Combined, these different effects could impact the measured sensory distances among products and the descriptions of such products. Yet, there is a lack of literature on this topic. And indeed, at the DA panel scale it is impossible to measure each effect separately but the outcome observed is the result of their combination.

The main goal of the present study is to determine if the addition of extra products modifies the outcome of sensory descriptive analysis on a small target set, and if so to what extent. Furthermore, this study aims at providing a strategy for choosing or engineering products to be added. Specifically, two factors were studied: the number of products added to the target set and the sensory proximity between the products added and the target set. The responses studied here were the attributes generated by the panel, the structure of the target dataset (distances among target products) as well as the attributes used to describe each product. All experiments were replicated on ketchup and lemonade to check if conclusions are product-dependent.

2. Materials and methods

2.1. Experimental design

This study was conducted in three parts, summarized in Fig. 1. Three different experiments were carried out. Experiment 1 tested whether training a panel on a larger set of products had an effect on the list of attributes, by training panelists on 8 products instead of the 4 in the two other experiments. Experiment 2 was designed to test whether the type of products added had an effect on the structure of the description of the target product set by adding two extra products, either similar to or different from the target products. Experiment 3 aimed at evaluating the effect of the number of products added testing two different conditions (four and six extra products). Each study consisted in a descriptive analysis protocol (Lawless & Heymann, 2010), using a different panel for each experiment, leading to a total of six different panels (two for each experiment, one using ketchup and the other one using lemonade). Each study was replicated on ketchup and lemonade to assess whether the results were reproducible on different product categories

2.2. Products

2.2.1. Product selection

Four commercially available ketchups and lemonades were used as the target product set (Tables 1a and 1b). The four target products were chosen to represent the diversity of the product category. In Experiment 1, four products were added to train panelists on eight products. They were chosen so that two were rather similar to the target products and two were different. Then, products were chosen for Experiments 2 and 3, according to the type and number of products needed. Tables 1a and 1b summarize the products used in each experiment. In Experiment 2, the same four added products as in Experiment 1 were reused, split into two conditions: one with two added samples, rather similar to the target and a second condition with two added samples, different from the target set. In Experiment 3, all four added training products from Experiment 1 were added for evaluation on 8 products (Fig. 1). Then, another two products were added for the condition with six extra products. One of these was chosen as relatively close to others while the other one was chosen to be rather different from the target products, to balance any effect of the type of products. These products were chosen based on the results of a sorting task and previous descriptive analysis data (Lestringant et al., 2017b).

2.2.2. Sample preparation

Details regarding the different types of lemonades (shelf-stable, powder or refrigerated) can be found in Table 1b. Powdered lemonades that required dilution in water were prepared according to directions on the label. All samples were presented at room temperature in 2-oz black plastic cups labeled with 3-digit random numbers. Serving sizes were about 10 g for ketchup and about 40 mL for lemonade.

2.3. Descriptive analysis protocol

Panelists for descriptive analysis were recruited based on availability and motivation. They were not told the purpose of the study. They had varying levels of experience in sensory descriptive analysis. In Experiment 1, there were 7 judges in the ketchup panel (4 females and 3 males, 23–66 (average 45) years old) and 8 judges in the lemonade panel (4 females and 4 males, 18–50 (average 27) years old). The panels in Experiment 2 consisted of 11 panelists (9 females and 2 males, 19–41 (average 31) years old) for ketchup, and 12 panelists (6 females and 6 males, 19–34 (average 23) years old) for lemonade. In Experiment 3, the ketchup panel consisted of 10 panelists (6 females and 4 males, 21–37 (average 24) years old) and the lemonade panel was composed of 11 panelists (7 females and 4 males, 21–44 (average 27) years old). Only one panelist tasting ketchup performed both Experiments 2 and 3. These were separated in time by three months and, like other panelists, they were not aware of the goal of the experiment. Each of the six descriptive panels followed the same training and evaluation protocol. There were six training sessions of 1 h each, during which panelists generated descriptors to describe the aroma, taste and mouthfeel of the products. Panelists discussed the relevance of each descriptor based on reference standards (listed in Appendix A–F). All references were revised until they were approved consensually by panelists. Finally, judges evaluated all the descriptors on 15-cm unstructured scales anchored at both ends with "None" and "Very intense", except for attributes "Thick" ("Very thin/Very thick"), "Grainy" ("Very smooth/Very grainy") for ketchup and "Viscous" (Very watery/Very viscous) for