Contents lists available at ScienceDirect

Food Quality and Preference

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

Optimized Descriptive Profile: How many judges are necessary?

CrossMark

Rita de Cássia dos Santos Navarro da Silva [a,*], Valéria Paula Rodrigues Minim [a], Alexandre Navarro da Silva [b], Luiz Alexandre Peternelli [c], Luis Antônio Minim [a]

[a] Departamento de Tecnologia de Alimentos, Universidade Federal de Viçosa (UFV), 36570-000 Viçosa, Minas Gerais, Brazil
[b] Departamento de Engenharia de Produção e Mecânica, Universidade Federal de Viçosa (UFV), 36570-000 Viçosa, Minas Gerais, Brazil
[c] Departamento de Estatística, Universidade Federal de Viçosa (UFV), 36570-000 Viçosa, Minas Gerais, Brazil

ARTICLE INFO

ABSTRACT

The cost associated with descriptive sensory tests can be derived primarily from two sources: (i) execution time of the test and (ii) the number of judges participating. The Optimized Descriptive Profile (ODP) technique is a new methodology that proposes to reduce test time through an optimized sensory evaluation protocol. The objective of this study was to determine the optimal number of judges for descriptive sensory evaluation using the ODP, so as to show that the technique presents a reduction in time, money and efforts for conducting the methodology, and also in relation to the number of participants making up the panel. The study to determine the optimal number of judges was conducted using the data re-sampling technique for a panel original composed of 26 judges, by means of computer simulation. Data from the complete panel considering 10,000 sub-groups was re-sampled with replacement. The criteria for determining the ideal number of judges were: (i) acquisition of an experimental error less than or equal to the error verified in the reference methodology (Conventional Profile), (ii) obtaining interaction between sample and judges, in terms of size and stability, similar to interaction obtained by complete panel, (iii) concordance rate among products, using paired comparison (sample discrimination), similar to the full panel and (iv) minimal loss of information in the sensory map. The criterion for magnitude of the experimental error estimate showed to be the most robust measure for determination of the number of judges necessary for the ODP technique. Because this technique requires low levels of training of the judges, evaluation of these criteria is extremely important since a larger residual random variation can usually be observed. The criteria for magnitude of the experimental error, interaction between samples and judges effect and concordance rate in paired comparisons were met when sixteen evaluators was used.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The cost associated with descriptive sensory evaluations increases with the number of participating judges. Therefore, determination of the ideal number of evaluators is of utmost importance. According to Heymann, Machado, Torri, and Robinson (2012), it is obvious that training a smaller number of judges requires less time, cost and effort, but this may result in a "false savings" due to the possibility of obtaining "poor" data. Thus, the challenge is to determine the optimal number of judges needed for descriptive assessments that allows for reducing the size of the panel, but without information losses on the sensory profile of the foods.

The recommended optimal number of judges composing a panel is not very clear in literature. Different recommendations are encountered depending on the technique used, for example, six judges for the Flavor Profile (Cairncross & Sjostrom, 1950), ten judges for the Texture Profile (Brandt, Skinner, & Coleman, 1963) and ten to twelve judges for the Quantitative Descriptive Analysis (Stone & Sidel, 1985). However, the criteria for determining the number of judges needed are not shown.

Calculation of the number of judges in descriptive sensory testing has been little explored in literature. Some studies were conducted to determine the optimal number of judges considering generic methodologies, such as the "Conventional Profile" or "Descriptive Analysis". In most studies, reduction in the number of judges making up the panel was addressed by means of re-sampling data obtained by larger panels (Gacula & Rutenbeck, 2006; Heymann et al., 2012; King et al., 1995; Pagès & Périnel, 2003 and Silva, Minim, Silva, & Minim, 2014).

* Corresponding author. Tel.: +55 31 38993810.
E-mail addresses: rita.navarro@ufv.br (R.C.S.N. Silva), vprm@ufv.br (V.P.R. Minim), alexandre.navarro@ufv.br (A.N. Silva), peternelli@ufv.br (L.A. Peternelli), lminim@ufv.br (L.A. Minim).

King, Arents, and Moreau (1995) conducted a sensory description of ice cream samples by performing an evaluation with a panel of 20 judges (full panel). Data from the full panel was re-sampled by 20 smaller panels, consisting of 3–13 judges. For the new panels formed the significance of the treatment effects was evaluated for each of the sensory attributes by Analysis of Variance (ANOVA). The full panel presented greater explanation of the variation of the treatment effect, where 76% of the evaluated attributes were significant ($p < 0.05$) in the ANOVA. When the panel was reduced to half ($N = 10$), 67% of the attributes presented significance. Further reducing the panel to one quarter ($N = 5$) made only 34% of the attributes significant. The authors concluded that the reduction in number of judges in the panel resulted in significant loss of information regarding the effect of the treatments.

In the study performed by Pagès and Périnel (2003), a sensory description of eight samples of carbonated mineral water was evaluated by a panel of sixteen judges. Data obtained by the full panel was removed from the data set, two at a time, until reaching the minimum number of two assessments by the panel. For the sub-panels formed, the magnitude of the F-ratio and sensory map obtained by the PCA (Principal Components Analysis) were evaluated. No difference between the panels was observed considering these criteria.

Gacula and Rutenbeck (2006) determined the number of judges for descriptive sensory tests by computer simulation using experimental data. A panel of six trained judges proceeded to evaluate the samples in obtaining the data. In the simulation two experimental measurements were considered: difference to be detected between the means (d') and the variability of the experiment (Root Mean Square Error – RMSE). A minimal number of five judges was determined to make up the sensory panel.

Heymann et al. (2012) conducted a study on the number of judges for descriptive tests by re-sampling the original data. Data from three studies on sensory characterization of red wines was used, which included 14–22 judges. Data obtained by the complete panels was re-sampled by panels with 4, 6, 8, 12 and 14 judges. The new panels of judges formed were evaluated regarding significance (p-value) of attribute descriptors by the ANOVA and the sensory maps obtained by multifactor analysis. The results showed that at least eight judges are needed for the sensory panel.

The study by Silva et al. (2014) calculated the optimal number of judges for "Power analysis and Sample size". Three levels of probability were determined for the Type I and Type II errors and the difference to be detected in the experiment (d'). The standard deviation values of the experimental error were determined based on data from literature. A total of 574 values of the root mean square error (RMSE) were obtained from previous studies. Data from literature was adjusted for a known probability distribution, using 5% of this distribution in calculation of the number of judges. The required numbers of assessments in the descriptive tests were calculated, considering these different experimental conditions, totaling 135 scenarios.

The Optimized Descriptive Profile (ODP) methodology presents no previous studies concerning determination of the number of judges needed in the sensory panel. It was recently proposed as a descriptive method, and therefore there are few studies on this new sensory technique (Silva et al., 2012, 2013). The technique uses an optimized evaluation protocol, presenting a comparative evaluation between the samples, followed by a quantitative evaluation using an interval scale. Because the technique recommends the participation of judges with a low degree of training and the evaluation protocol of the foods is different, it is necessary that a particular study of this method is performed to determine the optimal number of judges.

This study sought to determine the optimal number of judges for the sensory descriptive analysis of foods using the Optimized Descriptive Profile (ODP), in order to permit that the technique shows a reduction of time, money and efforts for conducting the methodology and also in relation to the number of panel members.

2. Materials and methods

Determination of the required number of judges for the ODP technique was performed using the re-sampling technique for data obtained by an original panel consisting of 26 judges by means of computer simulation. Data from the full panel was re-sampled considering 10,000 iterations with replacement. The experimental data was obtained by means of sensory characterization using the ODP technique for two food matrices: strawberry-flavored yogurt (Experiment A) and chocolate (Experiment B). The criteria for determining the optimal number of judges were: (i) obtaining an experimental error less than or equal to the error found for the reference methodology (Conventional Profile), (ii) obtaining interaction between samples and judges, in terms of size and stability, similar to the interaction obtained by the full panel, (iii) concordance rate among products, using paired comparison (sample discrimination), similar to the complete panel and (iv) minimal loss of information in the sensory map.

2.1. Stimulus

Two types of food matrices were used (yogurt and chocolate) in the sensory characterization. The formulations were defined based on preliminary triangular tests, in which the samples presented a small magnitude of difference ($p < 0.10$) in the sensory characteristics, showing proportion of distinguishers (Pd) minor than 0.29 in the guessing model, equivalent to d' equal to 1.6 in Thurstonian model. The probability of error Type II was established at 0.10.

2.1.1. Experiment A
Five strawberry flavored yogurt formulations were utilized. A commercial brand yogurt was used for preparation of the formulas. Different concentrations of milk, sugar, powdered milk and pink dye were added at different concentrations (Table 1).

2.1.2. Experiment B
Four chocolate formulations were used. Chocolate formulations were prepared with three different chocolate types from the same brand and each unit measured 29 mm in diameter and 20 mm in height. In preparation of the formulations different mixtures of milk chocolate, semisweet chocolate and bitter were used. The amounts of each type of chocolate used in the process are described in Table 2.

2.2. Procedure

Sensory evaluation of the test-formulations (yogurt and chocolate) was performed using the evaluation protocol of the Optimized Descriptive Profile technique, (Silva et al., 2012, 2013). Thus, two panels of 26 judges participated in evaluations of the strength attribute descriptors (Tables 2 and 3). One panel of 26 judges performed a sensory evaluation of the five yogurt formulations and another panel of 26 judges evaluated the sensory characteristics of four chocolate formulations.

For ODP technique, the judges were recruited by structured questionnaires and pre-selected by difference tests (e.g., triangular tests). They also defined the sensory attributes for descriptive evaluation of the samples and the reference materials for each attribute. The judges evaluated the products in relation to the sensory attributes using a 9 cm unstructured rating scale, with three repetitions, according to the ODP protocol.