Food Quality and Preference 49 (2016) 106–118

Contents lists available at ScienceDirect

Food Quality and Preference

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

When satiety evaluation is inspired by sensory analysis: A new approach

Aurélie Lesdéma [a,b,*], Agnès Marsset-Baglieri [b], Liliane Talbot [a], Agathe Arlotti [a], Julien Delarue [c], Gilles Fromentin [b], Marie-Christine Marcuz [a], Sophie Vinoy [a]

[a] Mondelēz International Biscuit RD&Q, 91400 Saclay, France
[b] AgroParisTech, INRA, CNRH-IdF, UMR914 Nutrition Physiology and Ingestive Behaviour, F-75005 Paris, France
[c] AgroParisTech-UMR1145, Sensory Laboratory, Massy, France

ARTICLE INFO

Article history:
Received 11 June 2013
Received in revised form 11 December 2014
Accepted 15 November 2015
Available online 17 November 2015

Keywords:
Training
Satiety
Visual analog scale
Sensory panelists
Repeatability

ABSTRACT

The current approach for evaluating feelings of satiety using visual analog scales (VAS) is well developed. Although widely validated, there are certain limitations inherent to VAS, such as the difficulty of understanding and using them or the limited introspection naïve consumers have for evaluating appetite sensations. The hypothesis of the new approach presented here is that selecting sensory panelists used to evaluating feelings about foods, then giving them specific training followed by application to foods in the assessment of appetite feelings provides more discriminative results.

Indeed, one goal of our work is to test whether such training increases panel performance and, in particular, the homogeneity of the panel, leading to greater power and consequently allowing smaller panels which would prove useful when a large number of products must be tested. Eighteen sensory panelists were trained to evaluate appetite feelings using VAS. Similarly to the methodology used with sensory panelists before conducting descriptive analysis (Lawless & Heymann, 2010), this training was dedicated to: (1) understanding the vocabulary to define appetite and (2) manipulating the scales (3) using the scales by evaluating appetite sensations on a large variety of carbohydrate-rich products. During the third part of the training (4 sessions) and during 6 subsequent 170-min sessions, fasting panelists had to consume 1047 kJ of a single food product each session and to fill out VAS every 30 min. Seven different products were evaluated. Among them a reference product was tested 4 times. This methodology was applied to a second panel in order to evaluate the reproducibility of the method. In addition, both a group of naïve consumers and the trained sensory panelists evaluated 4 biscuit prototypes so that their results could be compared. 4 VAS questions were used to calculate an average appetite score. No difference in appetite score was obtained between the repetitions of the method with the reference products, which showed very similar appetite scores throughout time. Rankings based on satiety feelings were obtained for both a large variety of carbohydrate-rich foods and biscuit products. In addition, this methodology was reproducible when applied to another group of panelists. This approach also showed stronger discrimination of perceived satiety power by the panelists than by naïve consumers. Panelists were successfully trained as satiety experts capable of building a reproducible and discriminatory method which can be used regularly. Such training may help focus the evaluation of appetite feelings on physiological cues. This quick training provides a relevant way of screening foods with potential interest for further investigation on clinical trials of food.

1. Introduction

Several methods have been developed and are commonly used to evaluate food intake, appetite sensations and/or the underlying mechanisms (Blundell & Bellisle, 2013). These methods include objective measurements such as *ad libitum* food intake at the next lunch following preload consumption, the delay between two meals, keeping a food diary (Griffioen-Roose et al., 2012; Higgs, Williamson, & Attwood, 2008; Westerterp-Plantenga et al., 1998) or measurement of biological satiety markers (Blundell et al., 2010). Satiety can also be measured subjectively by using visual analog scales (VAS). These scales are commonly used and validated to measure appetite sensations (Flint, Raben, Blundell, & Astrup, 2000; Stubbs et al., 2000). This tool is known to have good repeat reliability between groups, which means that the same and

* Corresponding author at: Mondelēz International Biscuit RD&Q, 6 rue René Razel, 91400 Saclay, France.
E-mail address: aurelie.lesdema@mdlz.com (A. Lesdéma).

http://dx.doi.org/10.1016/j.foodqual.2015.11.004
0950-3293/

compatible results can be produced in other clinical trials (Blundell et al., 2010; Flint et al., 2000; Merrill, Kramer, Cardello, & Schutz, 2002; Stratton et al., 1998; Stubbs et al., 2000). However, the VAS approach shows certain limitations. Inter-individual differences which can appear in the use and understanding of visual analog scales (Raben, Tagliabue, & Astrup, 1995), and the high variability in experimental designs (delay, type and energy of preloads, etc.) have to be considered (Leidy et al., 2015; Ortinau, Hoertel, Douglas, & Leidy, 2014; Pentikainen et al., 2014). The complexity of hunger and fullness sensations of consumers was highlighted by Murray and Vickers through the use of focus groups (Murray & Vickers, 2009). They concluded that sensations of mental hunger and physical fullness overlapped, which provided evidence that the overall constructs of hunger and fullness may not be simple, polar opposites. Reducing inter-individual variability could improve the precision of satiety assessment. In the context of satiety claims, according to the European Food Safety Authority (EFSA), it is possible to make claims on changes in appetite ratings. However 'the beneficial physiological effect of changing appetite ratings depends on the context of the claim'. Hence, evidence on changes in appetite ratings alone may not be sufficient for the scientific demonstration of the claim (EFSA Panel on Dietetic Products, 2012; Halford & Harrold, 2012). Indeed, in many cases satiety is evaluated with both subjective and objective measurements (Allison, 2009; Blundell et al., 2010). It could thus be interesting to find a methodology not intended as a replacement of the traditional one but which includes measurement of food intake while making quicker and efficient screening of a lot of products according to their satiety power possible.

In this context, our aim was to develop a new methodology relying on a specific population of sensory experts, herein referred to as 'panel 1', and to train them at evaluating their appetite sensations. We indeed hypothesized that we could take advantage of such subjects' pre-existing expertise in the description of their sensations. We sought to test whether the training would increase panel performance and, in particular, the homogeneity of the panel, leading to greater power and consequently allowing smaller panels which would be useful when a large number of products must be tested. Therefore, the subject's ability to discriminate such products according to their satiety power should be improved. Based on our hypothesis, training people should help to reduce differences originating in inter-individual variability such as a gender, age, food behavior, lack of understanding or difficulties in evaluating the intensity of hunger feelings along a unipolar unstructured line. Then, we compared the results produced by trained panelists with the appetite scores produced by naïve consumers by using four different types of biscuits.

The strength and reproducibility of this training was confirmed by applying the same methodology with a second panel herein referred to as 'confirming panel'.

2. Materials and methods

2.1. Panelists

Eighteen volunteers were recruited for panel 1 as well as 18 for the confirming panel. For panel 1 (one man and seventeen women), the average age was 50 ± 0.2 years and the mean BMI was 22.4 ± 0.4 kg/m^2. For the confirming panel (five men and thirteen women), the average age was 50.0 ± 0.5 years and the mean BMI was 22.7 ± 0.1 kg/m^2. For $p = 0.05$, $\sigma = 12$ mm, $\Delta = 10$ mm for the appetite score and a power level of 0.9, we needed about 15 subjects (Jmp 10 software and (Whitley & Ball, 2002)). Therefore 18 subjects were recruited in order to retain enough subjects once the potential outliers had been removed. All panelists were members of sensory expert panels and are thus herein referred to as panelists or judges. Their fields of expertise were diverse. Some were specialized in the sensory evaluation of food products such as drinks, dairy products or cereal products, while others had extensive training in the evaluation of flavoring or cosmetics. Therefore, the main criterion was their sensorial expertise ability. As these sensory panelists are used to evaluating sensations with different scales, we thought they could adapt to VAS ratings more quickly than naïve consumers. We expected them to be quicker at learning than people who are not used to sensory scales or to focusing on their sensations. They were non-smokers, with no weight variation of over 2 kg for the 2 months preceding the study, presented no food allergies or dislikes regarding the foods proposed for the test meals and were used to eating breakfast every day. Participants had previously filled out a questionnaire to make sure that they had no aversion regarding biscuits. In order to characterize their food habits, they were asked to complete the Three-Factors Eating Questionnaire, which was used to evaluate their cognitive restraint (TFEQ-R). Their average TFEQ-R score was 7.2 ± 0.3 for panel 1 and 7.9 ± 0.2 for the confirming panel, which can be considered as low levels of cognitive restraint based on the limits commonly used (Stunkard & Messick, 1985).

2.2. Naïve consumers

Fifty-six women were recruited by advertisements listed on an Internet website specialized in volunteer recruitment, messages sent to electronic mailing lists and flyers posted around the laboratory and universities in Paris. All subjects were in a healthy state with similar criteria of exclusion to those used for the panelists. All volunteers recruited signed a consent form as well as an information package. The subjects received financial compensation for their participation, and the whole nature of the study was revealed to them during a debriefing session at the end of the study. The final sample was thus composed of 56 women with a mean BMI of 21.8 ± 0.2 kg/m^2, and a mean age of 24.5 ± 0.6 years. Their average TFEQ-R score was 4.0 ± 0.3, which can be considered as low levels of cognitive restraint based on the limits commonly used (Stunkard & Messick, 1985). The study was approved by the Human Research Ethics Committee of Aulnay. For $p = 0.05$, $\sigma = 20$ mm, $\Delta = 10$ mm for the appetite score and a power level of 0.9, we needed more than 50 subjects (Jump 10 software and (Whitley & Ball, 2002)). Therefore, enough subjects were recruited to assure a suitable number of subjects once the potential drop-out had been removed and to fit with the Latin square.

2.3. Training

The training of both panel 1 and the confirming panel was divided into 3 steps: (1) understanding the vocabulary to define appetite (2) manipulating the scales to help panel members evaluate the intensity of their hunger feelings on unipolar unstructured line scales (3) using the scales by evaluating appetite sensations on a large variety of carbohydrate-rich products. The training involved 5 sessions: one dedicated to the 2 first steps and 4 sessions dedicated to the third step. This method was duplicated with the confirming panel afterwards to confirm the initial results.

The training concerned mastering the vocabulary used in VAS to describe appetite sensations and design a satiety lexicon. The aim was to teach the panelists how to specifically recognize and evaluate the intensity of their hunger sensations (hunger, fullness, desire to eat and prospective consumption) and to report their sensations in a "satiety lexicon," to be used during the other sessions. One session of 3 h was dedicated to this part of the training. Methodologies commonly used in conventional descriptive sensory analysis were applied (Lawless & Heymann, 2010). This involved the 2 first