CrossMark

Hedonic product optimisation: CATA questions as alternatives to JAR scales

Gastón Ares [a], Juliana Cunha de Andrade [b], Lucía Antúnez [a], Florencia Alcaire [a], Marianne Swaney-Stueve [c], Sheri Gordon [c,d], Sara R. Jaeger [e,*]

[a] Sensometrics & Consumer Science, Facultad de Química, Universidad de la República, By Pass de Rutas 8 y 101 s/n, CP 91000 Pando, Canelones, Uruguay
[b] Federal University of Rio de Janeiro, Avenida Athos da Silveira Ramos, 149 – bloco A, CEP 21.941-909, Cidade Universitária, Rio de Janeiro, RJ, Brazil
[c] Sensory & Consumer Research Center, Kansas State University, 22201 W Innovation Drive, Olathe, KS 66061, USA
[d] Campbell Soup Company, Campbell Place, Camden, NJ, USA
[e] The New Zealand Institute for Plant & Food Research Limited, 120 Mt Albert Road, Private Bag 92169, Victoria Street West, Auckland, New Zealand

ARTICLE INFO

ABSTRACT

Methodologies that identify ways products differ from consumers' ideal are commonly used to guide innovation. In this research the use of CATA questions for this purpose was compared to JAR scales, which are well established for use in product optimization efforts. Two CATA variants were considered: (i) CATA questions including terms with hedonic-intensity connotations (e.g., *not enough sweet*, *much too sweet*), and (ii) CATA questions pertaining to both the tasted and the ideal products. In six consumer studies (n = 939), spanning multiple product categories and consumer populations (Uruguay, New Zealand and USA), it was found that CATA questions and JAR scales provided similar insights regarding the most relevant deviations from ideal. However, several differences were also identified. In particular, CATA questions tended to identify fewer deviations than did JAR scales, especially when terms with hedonic-intensity connotations were used and when differences between samples were small. This difference is likely linked to facets of the two methodologies: only applicable terms are selected when using CATA questions, whereas responses must be provided for each JAR scale included in a study. Besides, the fact that only the extreme anchors of the JAR scale were included in the CATA question could have encouraged consumers not to indicate deviations from the ideal. Penalty analysis, performed using Partial-Least Squares (PLS) regression identified several significant deviations from the ideal. While the two methodologies established the same main differences, JAR scales identified more significant deviations from the ideal than CATA questions. Although results confirmed the potential for the use of CATA questions in product optimization research, careful consideration of purpose of the research and attention to terms included in the CATA questions is recommended.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

New product development is a key activity for food companies to maintain growth, increase competitiveness and achieve long-term profitability (Costa & Jongen, 2006). Critical steps in the new product development cycle include the identification of prototypes that closely align to consumer preferences and obtaining direction for improving product performance (van Kleef, van Trijp, & Luning, 2006). To achieve this objective, information about consumers' perception of the sensory characteristics of the products is often concurrently collected with overall liking scores (Lawless & Heymann, 2010).

Just-about-right (JAR) scales are a well-established approach to identifying the optimum intensity of sensory attributes (Popper & Kroll, 2005). Using a 5-point bi-polar scale, consumers are asked to indicate whether the intensity of a sensory attribute is too strong, too weak or just-about-right (Popper, 2014). To provide directional information for improving product performance, JAR and hedonic responses are jointly analysed, typically using penalty analysis. The sensory attributes with the largest detrimental effect on consumer liking are identified, which can guide product reformulation (Lesniauskas & Carr, 2004).

Although JAR scales are a popular approach to gathering actionable information for product development, they are no panacea. According to Moskowitz (2001), JAR scales can be challenging for consumers because they require evaluating the perceived intensity of the sensory attribute, the ideal attribute intensity and the

difference between these perceived and ideal intensities simultaneously. In addition, JAR scales make consumers focus their attention on specific sensory attributes, which may increase awareness of how the product falls short from the ideal and ultimately change their hedonic perception of the sample (Epler, Chambers, & Kemp, 1998; Popper & Kroll, 2005; Popper, Rosenstock, Schraidt, & Kroll, 2004). For these reasons, alternative approaches to study consumers' perception of the sensory characteristics of products have been recently proposed.

Check-all-that-apply (CATA) questions are currently one of the most popular approaches for getting consumer-based sensory characterizations (Ares & Jaeger, 2015). In this methodology, consumers are presented with a list of sensory terms and are asked to select all those they consider applicable for describing the focal sample. This methodology has been reported to be easy and intuitive for consumers, providing valid and repeatable product information (Ares et al., 2015; Dooley, Lee, & Meullenet, 2010; Jaeger, Chheang, et al., 2013). Furthermore, CATA questions are not likely to influence consumers' hedonic responses (Jaeger & Ares, 2014; Jaeger, Giacalone, et al., 2013), probably because the task does not require consumers to strongly focus their attention on the list of terms (Krosnick, 1999; Strack, 1992). These features make CATA questions an interesting methodological choice for hedonic product optimization. Penalty analysis has been recently applied with data from CATA questions as a way to identify deviations from optimal and guide product improvement when hedonic scores are concurrently collected (Ares, Dauber et al., 2014; Meyners, Castura, & Carr, 2013).

The present research continues the exploration of CATA questions for use in product optimization tasks by: (1) considering two variants of CATA questions tailored for use in product improvement/optimization, and (2) contrasting these with JAR scales. The first CATA approach consisted of including in the CATA question terms with hedonic-intensity connotation similar to those used as anchors in JAR scales (e.g., *not enough sweetness*, *too much sweetness*). The second approach consisted of asking consumers to characterise both the tasted products and their ideal product using a CATA question, similarly to what is done with intensity scales in the ideal profile method (van Trijp, Punter, Mickartz, & Kruithof, 2007; Worch & Punter, 2014).

The new insights from these two approaches to using CATA questions to identify deviations from ideal and the comparative aspect relative to JAR scales will help practitioners to make informed decisions regarding product optimization tasks.

2. Materials and methods

Six consumer studies were conducted (Table 1). Studies 1 and 2 considered CATA questions including terms with hedonic-intensity connotation, and Studies 3–6 considered CATA questions involving the evaluation of the ideal product relative to JAR scales. The studies were carried out in three different countries (New Zealand, Uruguay and USA) and involved a range of product categories. In all studies a between-subjects experimental design was used to compare CATA questions and JAR scales.

2.1. Studies comparing CATA questions including terms with hedonic-intensity connotation and JAR scales

2.1.1. Participants

Studies 1 and 2 were conducted in Montevideo (Uruguay) and in Dallas and Olathe (USA), respectively, with a total of 508 consumers (Table 1). Participants were recruited from the consumer database of the research groups who led the studies. Participants gave informed consent and were compensated for their participation.

Consumers were aged between 18 and 60 years old. The percentage of female participants was 65% in Study 1 and 53% in Study 2. The consumer sample comprised varying household composition and income levels but was not representative of the populations of the cities where the studies were conducted. Participants were users of the focal products.

2.1.2. Samples

Six strawberry samples corresponding to different cultivars were included in Study 1. For each of the samples, one strawberry was presented to consumers in a closed odour-free plastic container labelled with three-digit random numbers, at room temperature.

In Study 2, four samples of Mexican sauces were considered, two market samples and two odour-free prototypes developed by an industrial food producer. For each sample, 57 g were served in Styrofoam bowls with 3 tortilla chips for scooping. Samples were coded with 3-digit random numbers and presented at approximately 40 °C.

2.1.3. Experimental procedure

In Study 1 consumers were randomly assigned to one of two experimental treatments: CATA questions or JAR scales. In Study 2, two consumer groups with similar socio-demographic characteristics and product usage frequency were recruited in two different cities (Dallas and Olathe). Each group was assigned to a different experimental treatment: consumers in Dallas evaluated samples using CATA questions, whereas consumers in Olathe used JAR scales.

The attributes considered in the studies were selected from previous consumer studies conducted with the same product categories and the researchers' familiarity with the product. In Study 1, four attributes were evaluated using 5-point JAR scales: *sweetness, acidity, colour intensity,* and *firmness,* (1 = Not enough, 3 = Just about right, 5 = Much too). The CATA question comprised 12 terms with intensity/hedonic connotation, which were identical to the anchors of the JAR scales (Table 1). In Study 2, ten attributes were evaluated using 5-point JAR scales: *light/dark, colour 1, colour 2, smooth/chunky, ingredient 1, flavour 1, flavour 2, spicy, tart* and *salty*. Consumers were asked to try each of the samples and to describe it using a 5-point scale (1 = Not enough, 3 = Just about right, 5 = Much too).[1] The CATA question comprised a total 32 terms with intensity/hedonic connotation, which were similar to the ones included as anchors in the JAR scale (Table 1). For example, the JAR attribute *light/dark* was represented by four CATA terms: *too light, light, dark* and *too dark*. In this study, in contrast to Study 1, the just-about-right term phrasing was not used. For example, the JAR attribute *ingredient 1* was expressed in the following three CATA terms: *not enough ingredient 1, ingredient 1* and *too much ingredient 1*.

To facilitate penalty analysis, overall liking data were collected in both studies using a 9-point hedonic scale (1 = 'dislike extremely', 9 = 'like extremely'). In Study 2, appearance and flavour liking data were also collected but data are not analysed here. For each sample, hedonic data were always collected prior to CATA questions or JAR scales.

Samples were presented in sequential monadic order following a Williams' Latin square design. Following standard practice, the presentation order of the JAR scales was fixed across participants and samples, whereas presentation order of the terms of the CATA

---

[1] For reasons of commercial confidentiality, the exact wording of some of the attributes cannot be presented