Food Quality and Preference 35 (2014) 1–5

Contents lists available at ScienceDirect

Food Quality and Preference

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

Short Communication

Lack of evidence that concurrent sensory product characterisation using CATA questions bias hedonic scores

Sara R. Jaeger [a,*], Gastón Ares [b]

[a] The New Zealand Institute for Plant & Food Research Limited, 120 Mt Albert Road, Private Bag 92169, Victoria Street West, Auckland, New Zealand
[b] Departamento de Ciencia y Tecnología de Alimentos, Facultad de Química, Universidad de la República, Gral. Flores 2124, C.P. 11800 Montevideo, Uruguay

ARTICLE INFO

Article history:
Received 22 October 2013
Received in revised form 1 January 2014
Accepted 6 January 2014
Available online 13 January 2014

Keywords:
Consumer research
Sensory properties
Research context
Hedonic scale
CATA

ABSTRACT

In response to the growing use of consumers for sensory product characterisation, methodological research contributing to development of best practise guidelines is ongoing. We focus here on concurrent elicitation of hedonic and sensory product characterisation by check-all-that-apply (CATA) questions. Jaeger et al. (2013b) reported that CATA questions only caused weak and transient bias of co-elicited hedonic scores. In the current research six studies were conducted, in which more than 700 consumers took part. Five product categories were tested (rice crackers, lite bread, cheese, kiwifruit, black currant drinks) with 4–7 samples per study. In none of these studies was evidence obtained suggesting bias of hedonic scores and it is now possible to conclude with greater certainty that co-elicitation of hedonic scores and product attribute information using CATA questions is unlikely to bias hedonic scores. A second result of the current research was that the use of designs that rotate presentation order of CATA terms was not associated with hedonic bias, and neither was the use of the forced Yes–No CATA question format. In future research, in light of a strong dominance of positive CATA terms used in these studies, we recommend studying more thoroughly the influence of positive/negative/neutral words in CATA lists as a possible source of hedonic bias. An exploratory component to this research suggested that consumers perceive the concurrent elicitation of hedonic and CATA responses as easy, but that too many samples may make the task tedious.

1. Introduction

Methodological research into the use of consumers for sensory attribute characterisation has in part been prompted by the growth of such practice in academia and industry (Varela & Ares, 2012). This paper continues work by Ares, Jaeger and colleagues regarding the use of check-all-that-apply (CATA) questions for sensory product characterisation by consumers, which aims to generate guidelines for best practise (e.g., Ares & Jaeger, 2013; Ares, Tárrega, Izquierdo, & Jaeger, 2014; Ares et al., 2013; Jaeger et al., 2013a; Jaeger et al., 2013b).

We focus here on concurrent elicitation of hedonic and attribute diagnostic information using CATA. Several authors have reported that including questions about specific sensory characteristics can be a source of bias on hedonic scores (Prescott, Lee, & Kim, 2011; Stone & Sidel, 2004). CATA questions have been claimed to have a potential smaller effect than other attribute-based question types such as just-about-right or intensity scales (Adams, Williams, Lancaster, & Foley, 2007). Jaeger et al. (2013b) reported that CATA questions only caused weak and transient bias of co-elicited hedonic scores across a range of product categories (beer, fresh fruit, tea, flavoured water, crackers, savoury dips). This effect was independent on whether samples, on average were moderately liked or moderately disliked, and replicated when samples were assessed partially by the sense of smell only or via full product assessment (appearance, aroma, flavour, taste, aftertaste, mouth feel). On this basis the authors concluded that co-elicitation of hedonic scores and product attribute information using CATA questions may bias the hedonic scores, but not that it certainly will do so. This research seeks to confirm the robustness of these findings by using a more diverse set of product categories, as well as product sets that include multiple products.

Further we extend Jaeger et al. (2013b) by investigating whether recommendations to minimise saliency bias in CATA responding, through balancing presentation order of CATA terms (across or within participants) and/or forced Yes–No answering to all terms in a CATA question is associated with a higher risk that concurrent use of sensory CATA questions in hedonic tests will lead to bias of hedonic scores. Both these CATA question formats are developed with a view to encourage cognitive attention by participants to the CATA question (Krosnick, 1991, 1999). Yet, when used conjointly with hedonic scaling where CATA questions have been suggested as superior to other attribute questions because they

* Corresponding author. Tel.: +64 99257000; fax: +64 99257001.
E-mail address: sara.jaeger@plantandfood.co.nz (S.R. Jaeger).

0950-3293/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodqual.2014.01.001

do not encourage deep processing (Strack, 1992; Sudman & Bradburn, 1992; Rasinski Mingay & Bradburn, 1994), this practise has the potential to be detrimental in the sense of being associated with bias of hedonic scores.

2. Materials and methods

A total of six studies were conducted, in which more than 700 consumers took part. Five product categories were tested (rice crackers, lite bread, cheese, kiwifruit, black currant drinks) with 4–7 samples per study. Between-subjects designs were used in all studies.

2.1. Participants

A total of 712 consumers participated in the 6 studies. The number of consumers in each study ranged from 102 to 181. Participants who completed Study 3 also completed Study 6. Participants attended research sessions at the Plant & Food Research Sensory Facility in Auckland. Participants were registered on a database maintained by a professional recruitment firm and were screened in accordance with eligibility criteria for each of the studies. Participants gave voluntary consent to participate and were compensated in cash.

All studies were part of multi-product research projects with test methods, product selection and recruitment criteria specified to match the composition of each project. With the exception of Study 3 (kiwifruit) it was not a recruitment criterion that participants be regular consumers of the tested products. However, at the stage of recruitment they confirmed willingness to eat the products to be included in the test.

All participants lived in the greater Auckland region (New Zealand) and were mainly of Caucasian ethnicity (Studies 2–6). Participants in Study 1 were of Chinese ancestry and recent immigrants to New Zealand (less than 3 years) and the study was conducted in Mandarin. All participants were aged 18–67 years old and the percentage of female participants ranged from 50% to 65%.

Participants represented households in diverse socio-economic strata. None of the consumer samples were representative of the general New Zealand population.

2.2. Samples

Five different product categories were tested (Table 1). All samples in Studies 1–4 were commercially available in New Zealand and purchased from local supermarkets. In Studies 5–6, samples of blackcurrant juice were used. Whole blackcurrant fruit from commercial and pre-commercial selections were sourced from a research orchard. Fruit from each cultivar were pressed into juice using a 20 × 17 cm basket fruit press (Ferrari, 11075), poured into 300 ml plastic (HDPE) bottles with plastic lids and blast frozen. Microbiological tests were carried out by AgriQuality on the extracted juice to ensure it was food safe for a period of 5 days storage at 4 °C. A day prior to consumer testing, juices were thawed and stored at 4 °C. A commercially available concentrate (Barker's) was included as a comparative sample. Only this sample and juice from one other commercially available cultivar were included in both Studies 5 and 6. The Brix level of each sample was adjusted to 9.4 g/100 g sugar by addition of sucrose before making a 1:14 dilution with filtered water.

In Studies 1–6, serving sizes were always sufficient to allow 2–3 bites/sips per sample and samples were always presented at room temperature.

2.3. Procedure

The procedure for data collection in Studies 1–6 was similar. Between-subjects experimental designs were always used, comparing responses from 2 or 3 experimental treatments. As detailed in Table 1, Experimental Treatment A was always 'Hedonic only', meaning that participants in this group only provided hedonic responses to the presented samples (9-pt labelled category scale; 1 = 'dislike extremely', 9 = 'like extremely'). In Experimental Treatments B and C participants always provided both hedonic and sensory CATA responses (Table 1). The studies differed in the way that the CATA question was implemented: presentation order of CATA terms and/or forced Yes/No answering (Table 1). Three presentation orders of the terms were considered: fixed order of terms for all participants [F], unique order of terms for each participant that is (near) balanced for presentation order [A], and different order of the terms for each product and each participant that is (near) balanced for presentation order [W] (see Table 1). In studies 1–4 participants were asked to check all the terms from the list that were considered applicable for describing samples, while in Studies 5 and 6 they were asked to indicate if each of the terms included in the list were applicable (Yes) or not applicable (No).

The CATA terms used in each study were based on previous research using these product categories and/or pilot work with staff from Plant & Food Research. The CATA lists contained 12–16 terms and covered multiple sensory modalities (appearance, aroma, flavour/taste, texture, after taste/mouth feel) (Table 2).

Participants attended research sessions in groups of 10–12 people. All data collection took place in standard sensory booths under white light. Temperature and air flow was regulated. All samples were presented in cups labelled with 3-digit random codes. Products were presented in accordance with experimental designs that were balanced for order and carry-over effects.

Considering that the same participants completed Study 3 and Study 6, participants allocated to Treatment A did this treatment in both studies. Other participants completed Treatment B or C in both studies or a combination of these two treatments. This allocation of participants to experimental treatments was used to retain participants in 'similar mindsets' (i.e., hedonic only or hedonic + CATA).

In Studies 4 and 5 participants answered two Likert questions immediately following completion of the study: (i) It was easy to answer the questions about these samples; and (ii) It was tedious to answer the questions about these samples. The labelled 7-point scale was anchored at 1 = 'disagree extremely' and 7 = 'agree extremely'.

For classification purposes participants' age and gender was recorded. In Studies 1 and 3–6, frequency of consumption of the focal product category was also obtained. In all instances differences between the participant profiles of the experimental treatment groups were non-significant ($p > 0.15$). Hence, it is possible to infer that differences between experimental treatments may be linked to differences in study protocol, as opposed to differences in group characteristics.

Data were always collected as part of sessions that featured multiple tasks including several product categories and research methods. Only data relevant to the aims of this research are presented here.

2.4. Statistical analyses

For each study, linear mixed modelling was performed to uncover significant differences in hedonic ratings across experimental treatments. Treatments, samples and their interaction were specified as fixed effects, whereas consumer (within experimental