Food Quality and Preference 42 (2015) 27–36

Contents lists available at ScienceDirect

Food Quality and Preference

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

Maximizing overall liking results in a superior product to minimizing deviations from ideal ratings: An optimization case study with coffee-flavored milk

Bangde Li [a,b], John E. Hayes [a,b], Gregory R. Ziegler [b,*]

[a] Sensory Evaluation Center, College of Agricultural Sciences, The Pennsylvania State University, University Park, PA, USA
[b] Department of Food Science, College of Agricultural Sciences, The Pennsylvania State University, University Park, PA, USA

ARTICLE INFO

Article history:
Received 8 June 2014
Received in revised form 12 January 2015
Accepted 12 January 2015
Available online 20 January 2015

Keywords:
Consumer behavior
JAR scaling
Ideal scaling
Product optimization
Dissatisfaction
Coffee

ABSTRACT

In just-about-right (JAR) scaling and ideal scaling, attribute delta (i.e., "Too Little" or "Too Much") reflects a subject's dissatisfaction level for an attribute relative to their hypothetical ideal. Dissatisfaction (attribute delta) is a different construct from consumer acceptability, operationalized as liking. Therefore, we hypothesized minimizing dissatisfaction and maximizing liking would yield different optimal formulations. The objective of this research was to compare product optimization strategies, i.e., maximizing liking vis-à-vis minimizing dissatisfaction.

Coffee-flavored dairy beverages ($n = 20$) were formulated using a fractional mixture design that constrained the proportions of coffee extract, milk, sucrose, and water. Participants ($n = 388$) were randomly assigned to one of three research conditions, where they evaluated 4 of the 20 samples using an incomplete block design. Samples were rated for overall *liking* and for intensity of the attributes *sweetness*, *milk flavor*, *thickness* and *coffee flavor*. Where appropriate, measures of overall product quality (*Ideal_Delta* and *JAR_Delta*) were calculated as the sum of the absolute values of the four attribute deltas. Optimal formulations were estimated by: (a) maximizing *liking*; (b) minimizing *Ideal_Delta*; or (c) minimizing *JAR_Delta*. A validation study was conducted to evaluate product optimization models.

Participants indicated a preference for a coffee-flavored dairy beverage with more coffee extract and less milk and sucrose in the dissatisfaction model compared to the formula obtained by maximizing *liking*. That is, when *liking* was optimized, participants generally liked a weaker, milkier and sweeter coffee-flavored dairy beverage. Predicted *liking* scores were validated in a subsequent experiment, and the optimal product formulated to maximize *liking* was significantly preferred to that formulated to minimize dissatisfaction by a paired preference test. These findings are consistent with the view that JAR and ideal scaling methods both suffer from attitudinal biases that are not present when liking is assessed. That is, consumers sincerely believe they want 'dark, rich, hearty' coffee when they do not. This paper also demonstrates the utility and efficiency of a lean experimental approach.

1. Introduction

Cold coffee flavored beverages continue to grow in popularity; between 2009 and Q1 2013, the share of cold-served coffee beverages of all coffee on food service menus in the US increased from 19% to 24% (Mintel., 2013). Notably, millennials are heavy consumers of these beverages compared to boomers: as 38% of those 18–24 drink iced coffee versus 11% of individuals aged 55–64 (Mintel, 2013). These beverages can be formulated with either water or milk as base (e.g., Petit & Sieffermann, 2007). For most Americans, consumption of dairy products falls well below recommendations in *The Dietary Guidelines for Americans* (Hayden, Dong, & Carlson, 2013). This is consistent with data showing fluid milk consumption among children and adolescents in the United States has been declining since 1977–1978 (Hayden et al., 2013; Sebastian, Goldman, Enns, & LaComb, 2010). As flavored milks are very popular in both children and adults (Kim, Lopetcharat, & Drake, 2013), the introduction of new flavored milks may help individuals reach recommended intake of nutrients like calcium and vitamin D (Kim et al., 2013; Nicklas, O'Neil, & Fulgoni, 2013).

Coffee flavor can be a positive factor for consumer acceptance of a coffee beverage (Li, Hayes, & Ziegler, 2014a). However, increasing

* Corresponding author at: Department of Food Science, Pennsylvania State University, 341 Food Science Building, University Park, PA 16802, USA.
E-mail address: grz1@psu.edu (G.R. Ziegler).

http://dx.doi.org/10.1016/j.foodqual.2015.01.011
0950-3293/

coffee flavor by adding more coffee extract also increases bitterness, and excessive bitterness typically reduces consumer acceptance (Harwood, Ziegler, & Hayes, 2012; Hayes, Sullivan, & Duffy, 2010; Lanier, Hayes, & Duffy, 2005; Moskowitz & Gofman, 2007). The amount of milk in a coffee beverage influences not only the appearance and the amount of milk flavor, but other attributes via physicochemical interactions. For example, the casein found in milk reduces the bitterness of coffee (Parat-Wilhelms et al., 2005). Therefore, a trade-off decision has to be made to reach an optimal formulation, which can be assisted using optimization techniques. Here, we optimized coffee flavored fluid milk (coffee milk) using two distinct approaches, and we describe the insights gained in the process.

Optimization is an important practice for product developers and sensory specialists (Ares, Varela, Rado, & Giménez, 2011; Dutcosky, Grossmann, Silva, & Welsch, 2006; Villegas, Tarrega, Carbonell, & Costell, 2010). Given intense competition in the market, the food industry is perpetually interested in optimization tools that are both rapid and cost effective. According, just-about-right (JAR) scales have gained popularity as an optimization technique because they are quick and easily executed by sensory analysts (Popper & Gibes, 2004; Rothman & Parker, 2009; Xiong & Meullenet, 2006). Operationally, optimization can be approached in two distinct ways: by maximizing overall acceptability (e.g., Deshpande, Chinnan, & McWatters, 2008; Youn & Chung, 2012) or by minimizing dissatisfaction.

Using JAR scaling, an attribute is evaluated for its appropriateness relative to some ideal (Rothman & Parker, 2009; Worch, Dooley, Meullenet, & Punter, 2010). This hypothetical ideal is designated "Just About Right" or "Just Right." Accordingly, a participant may indicate an attribute is "Too Little", "Too Much" or "Just About Right." Generally, when an attribute is "Too Little" or "Too Much", the product developer increases or decreases the amount of the ingredient that corresponds to the attribute. Thus, JAR scales are said to give directional guidance. This technique may be useful when developers have only a limited number of prototypes to evaluate (versus a designed experiment with a large number of prototypes covering a wide product space), but there has been little validation of this (Moskowitz, 2001), and textbooks (Stone & Sidel, 2004) recommended against replacing designed experiments with JAR scaling for product optimization. JAR scaling has been criticized for conflating the measurements of attribute intensity and consumer acceptability into one measurement scale (Moskowitz, Munoz, & Gacula, 2003). Additionally, JAR scales may suffer from other flaws that interfere with optimization, such as attitudinal biases unrelated to sensory properties, or a lack of attribute independence (Rothman & Parker, 2009).

As an alternative to JAR scaling, ideal scaling measures the perceived intensity of an attribute and the intensity of a hypothetical ideal separately (Gilbert, Young, Ball, & Murray, 1996; Rothman & Parker, 2009; van Trijp, Punter, Mickartz, & Kruithof, 2007; Worch, Le, Punter, & Pages, 2012). Unlike JAR scaling, where the ideal level (i.e., "Just About Right" or "Just Right") is fixed at the middle of the scale, ideal scaling allows a participant to place his or her hypothetical ideal anywhere along the line. The magnitude of "Too Little" or "Too Much" can then be estimated by the deviation (delta) between the perceived intensity and ideal intensity.

Using either ideal scaling or JAR scaling, the deviation from ideal (or delta), is a measure of dissatisfaction in regard to that specific attribute. The farther the attribute intensity deviates from the ideal level (i.e., the larger the delta), presumably the lower the product quality would be, and the more a consumer would be dissatisfied.

We believe it is important to distinguish between minimizing dissatisfaction, as is done when directional information from JAR or ideal scales are used to reformulate products, and maximizing liking, via designed experiments, as different a route to product optimization. Notably, in the Kano model, consumer dissatisfaction is not simply the opposite of satisfaction (Berger et al., 1993; Kano, Seraku, Takahashi, & Tsuji, 1984). Further, disparities in optimal levels for a single attribute obtained from JAR scaling and hedonic scores have been widely reported (Bower & Boyd, 2003; Daillant & Issanchou, 1991; Epler, Chambers, & Kemp, 1998; Shepherd, Smith, & Farleigh, 1989; van Trijp et al., 2007; Vickers, 1988). These differences may be greater when health-related attributes are rated.

Lovely and Meullenet (2009) compared four approaches to the optimization of a strawberry yogurt – external preference mapping (EPM), Euclidian distance ideal point mapping (EDIPM), landscape segment analysis (LSA), and JAR – and concluded that EPM, EDIPM and JAR produced equivalent results. They further concluded that JAR optimization was an acceptable alternative to more complicated preference mapping methods, and that LSA did not yield a superior product (greater liking). However, these techniques were not compared to a product formulated by simply maximizing liking, but only to the original product with the highest liking. They recommended further research in which direct comparisons of optimization strategies are made, and concluded that validation studies were a logical means to compare the efficacy of methods.

It is not certain that directional ratings (JAR or ideal) truly reflect a consumer's ability to know the ideal point and judge the magnitude of deviation from that ideal point (Moskowitz, 2001). Consumer data showing that maximal liking corresponds to minimal deviation from the ideal ("just right") would provide evidence for the validity of directional scales. Moskowitz (2001) addressed this question using a ½ replicate central composite design requiring 48 prototypes to optimize the visual appearance of pizza topping formulations, and concluded that creating a product for which the directionals are all simultaneously "on target" produced a highly acceptable, but not necessarily maximally acceptable product. Moskowitz (2001) suggested that the generalizability of these results should be tested with other food products and attributes, especially flavor.

Here we test the validity of directional ratings by comparing optimal formulations obtained by maximizing liking as compared to minimizing attribute deltas (dissatisfaction) for taste and texture of coffee-flavored dairy beverages. Furthermore, we take a lean experimental approach using a fractional, constrained mixture design for formulation with an incomplete block design for sensory analysis.

2. Materials and methods

This project comprised two studies, i.e., study I: product optimization, and study II: optimization validation. In study I, product optimization was conducted under three research conditions that differed in research ballot design (designated as Liking, Ideal and JAR). In study II, consumer overall liking and preference for two selected optimal formulations were evaluated separately. The method of product preparation was identical for both studies.

Informed consent was provided by the participants, and data were collected with the approval of the Penn State Office of Research Protections as exempt from IRB review under the wholesome foods exemption in 45 CFR 46.101(b)(6). Participants were compensated for their time.

2.1. Sample formulation and preparation

In study I, twenty coffee-flavored dairy beverages were formulated (Table 1) using eChip® software (Wilmington, DE) to create a fractional, mixture design with four constrained variables: coffee extract (3.0–5.0 wt%; Autocrat Sumatra 1397, Autocrat Natural