Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Food Science

Hedonic scaling: assumptions, contexts and frames of reference

Armand V Cardello

The role of assumptions, contexts and frames of reference in hedonic scaling are reviewed and discussed. The assumption that affect is bipolar is a long-held one, but recent studies in the food science literature have challenged this assumption through the use of separate scales for liking and disliking. These data and other theoretical and empirical arguments show that the bipolarity assumption is yet to be resolved. More certain are the effects of contextual variables in hedonic scaling. Recent research on the nature, relative degree of liking/disliking, and similarity of other stimuli to the target stimulus, the role of end-anchor labels on hedonic scales, and the impact of the framing of hedonic questions and their influence on the consumer's 'consideration set' are reviewed and discussed. The goal of the review is to inform food scientists about assumptions and contextual influences that operate in hedonic testing, so that these influences may be controlled and/or taken into account when using hedonic scales in consumer testing.

Address
U.S. Army Natick RD&E Center, Natick, MA 01760, USA

Corresponding author: Cardello, Armand V (avcardello@gmail.com)

Current Opinion in Food Science 2017, **15**:14–21

This review comes from a themed issue on **Sensory science and consumer perception**

Edited by **Herb Meiselman**

http://dx.doi.org/10.1016/j.cofs.2017.05.002

2214-7993/Published by Elsevier Ltd.

Introduction

Measurement is the assignment of numbers to objects according to rules, and these rules define different types of scales, for example, nominal, ordinal, interval, ratio [1]. Much research in psychophysics has been devoted to establishing these rules, creating different scale types, and comparing results obtained using these scales. The number of different scales that have been used to evaluate model tastants, food, smells, sounds, etc. is both impressive and staggering [2–5,6•,7•,8,9]. In addition, the number of studies comparing these different scale types in the recent sensory and consumer research literature continues unabated [10–20]. In comparing different scale types, the primary criteria for establishing that one scale is better than another is its mathematical precision and its ability to discriminate among products. However, these functional criteria must be traded off against 1) ease of use/comprehension of the scale by consumers, 2) efficiencies in data collection and 3) efficiencies in data analysis [5,21,22]. Thus, while a ratio scale, like magnitude estimation, may offer better mathematical properties than an ordinal scale, like ranking; if one only requires a rough order of preference, then ranking offers greater ease of use and efficiency of administration/data analysis, while still providing the level of discriminability necessary to make appropriate management decisions.

The efficiency argument is why the 9-pt hedonic scale, a category scale with only interval or, perhaps, ordinal properties, has been the primary method of hedonic scaling in food science, in spite of the continued development of more sophisticated techniques (see later sections for a discussion of these scales). Both the perceived lack of practical advantages of the latter scales and the continuing stream of studies that compare different scales in order to identify 'better' or 'worse' ones has led at least one prominent researcher to call for a de-emphasis of scaling research in sensory and consumer science [23].

In light of the above call, in the present paper I avoid pedantic analyses related to psychophysical scale types, levels of mathematical measurement, and 'better' or 'worse' scales. Instead, I address assumptions and general issues that impact all scale types. These include the assumption that affect is bivariate in nature, the effects of stimulus context, and the role of internal and external frames of reference in hedonic scaling.

The assumption of bipolarity of affect

One feature of the 9-pt hedonic scale [24,25] is that it is a univariate, bipolar scale. That is, it delineates a single dimension of affect ranging from one polar extreme (dislike extremely) to an opposite polar extreme (like extremely). In this way, it is similar to almost all other hedonic scales used in the food science literature, both older and newer. The precedent for this conceptualization of affect began with 19th Century psychologists, like Wundt and Fechner, and continued with the work of Beebe-Center [26,27], who used scales with 'pleasantness' and 'unpleasantness' on either side of a neutral point, and Young [28], who utilized scales that

varied from 'very great displeasure' to 'very great pleasure.' Almost all subsequent hedonic scales used in sensory, food and consumer science have followed this early lead.

However, in several recent papers, Kwak *et al.* [29••,30•,31••] have challenged this assumption of bipolarity through the use of either two unipolar scales (one for liking and one for disliking) [29••,30•] or a bivariate response grid (with liking on one dimension and disliking on the other) [31••]. In the first set of these studies [29••,30•], ratings obtained using either 9-pt or 7-pt unipolar hedonic scales, anchored with 'no opinion' on the low end and 'like/dislike extremely' on the high end, were compared to a 9-pt bipolar scale anchored at the poles with 'like/dislike extremely' and with 'neither like nor dislike' in the middle. Results showed that, while there was no difference in sensitivity to product differences between scales, the correlations between liking and disliking on the two unipolar scales were low (−0.01 to −0.74). In addition, they found differences in the internal preference maps generated using either the unipolar or bipolar scale data. The authors concluded that consumers 'tend to use more of the independent conceptualization than bipolar conceptualization' when judging hedonics.

Theoretical doubts about the bipolar nature of affect have been raised previously. Studies in the fields of attitude measurement, mood, emotions, and cognitive neuroscience have shown evidence that positive and negative affect exist simultaneously, although not at the extremes [32–44,45•,46]. In addition, other researchers have shown response differences between liking and disliking, such as response times for liking judgments that are significantly faster, more spontaneous, and less reflective of a controlled process of analysis than disliking judgments [47]. In addition, early research on the application of magnitude estimation to the measurement of hedonics showed that liking and disliking are not symmetrical (see Ref. [48], which also contains interesting data on the independent scaling of liking and disliking). Combining the above findings with those of Kwak *et al.*, one might well question whether the assumption of bipolarity for hedonic scaling is still valid and, if not, what does this mean for the validity of hedonic scales used in food science and psychology for the past hundred years?

First of all, the notion that the bipolarity assumption may be invalid is premature. A number of contrary studies have reported a high inverse relationship between positive and negative affect and a number of authors have suggested that the discrepancies in the literature may relate to differences in the definition of bipolarity [36,37,49–51] (see also Ref. [52] for a general discussion of valence in affect). That is, bipolarity can be defined as a strict inverse relationship between pleasure and displeasure, requiring a correlation between liking and disliking of −1.0. However, since empirical correlations range between −0.3 and −0.7 [37], this strict definition seems unlikely. An alternative and weaker definition merely excludes the two poles from existing simultaneously. In an analysis of this latter definition of bipolarity, Schimmack [37] cites evidence showing that, while pleasure and displeasure can exist simultaneously at low magnitudes, they are mutually exclusive at high magnitudes, and Russell and Carroll [51] conclude that 'reports of mixed feelings are invalid and are due to inappropriate items, random error, response styles, and misinterpretations.' Taking all the data together, the results present conflicting evidence, leading some researchers to argue that the controversy defines a paradox, that is, that affect is both bipolar and independent [50] (see also recent texts on the structure of affect and emotions [53] and on the measurement of affect in food-related research [54]).

From a practical point of view, it is still necessary to critically analyze the empirical results from Kwak *et al.* [29••,30•,31••]. First, the unipolar scales used in their studies [29••,30•] were earmarked by a peculiar use of 'no opinion' as the low end anchor, which does not establish a true zero point of affect and is not comparable to 'neither like nor dislike' on the bipolar scale to which it was compared. Similarly, their use of 7-pt and 9-pt unipolar scales may have produced better discrimination than the 5-pt like and 5-pt dislike 'halves' of their 9-pt bipolar scale simply by virtue of their greater length. Thirdly, the correlations between scales were calculated within foods, not across, creating a restriction of range that may account for the low correlations. Lastly, from the point of view of practicality and generalization to the literature, the 9-pt bipolar scale that they used was not the conventional 9-pt hedonic scale [25], since only 3 points on the scale were labeled.

In addition, other evidence has been reported that may further temper Kwak *et al.*'s findings. One piece of evidence comes from a multi-country study that compared two 5-pt unipolar scales of liking/disliking (anchored at the low end by 'no liking' and 'no disliking') with the traditional bipolar 9-pt hedonic scale for the evaluation of 21 food names [55]. This study showed high Pearson product-moment correlations between the mean ratings across stimuli for all scales in all countries (>0.93). The unipolar liking scales were highly positively correlated with the 9-pt hedonic scale, while the unipolar disliking scales were highly negatively correlated with both the latter scale and the unipolar liking scale, indicating that the scales are not independent of one another. In addition, in terms of sensitivity of the scales to stimulus differences, >85% of the time both scales (unipolar and bipolar) came to the same conclusion regarding significant differences ($p < 0.05$) among stimuli. In the other 15% of cases, the two unipolar scales or the bipolar scale revealed significant differences when the other did not. Although a