VARIATIONS IN HEDONIC RATINGS CHARACTERIZED BY SCALE POLARITY, SCALE TYPES AND ATTRIBUTES

Wannita Jirangrat[1], Kennet Carabante Ordonez[1]; Jing Wang[2], Jose Andres Herrera Corredor[3], Sujinda Sriwattana[4], and Witoon Prinyawiwatkul[1]

[1] Department of Food Science, Louisiana State University Agricultural Center, Baton Rouge, LA 70803, USA
[2] Department of Biostatistics, School of Public Health, Saint Louis University, St. Louis, MO, USA
[3] Colegio de Postgraduados, Campus Córdoba, Veracruz, México, C.P. 94946
[4] Sensory Evaluation and Consumer Testing Unit, Faculty of Agro-Industry, Chiang Mai University, Chiang Mai, Thailand 50100

ABSTRACT

The negative scale of the 9-point hedonic scale is not fully understood. When used to evaluate negative attributes (e.g., bitterness), this hedonic scale may yield poor results. We evaluated hedonic ratings as affected by scale types [9-point-categorical (CAT), line (LIN) and labeled-affective-magnitude (LAM)], polarity (unipolar vs. bipolar), and attributes (positive vs. negative). We compared sensitivity and confounding effects [contrast+panelist effects=CP] of positive- and negative-attribute ratings among 3 scales, and compared effects of uni-(negative-side only) vs. bipolar scales on negative-attribute ratings.

With bipolar scales, (1) consumers better differentiated negative-attribute ratings; (2) CP was higher for positive- than negative-attribute ratings [5.47 vs. 0.11, 12.41 vs. 0.09 and 82.66 vs. 0.23, respectively, for CAT, LIN and LAM; (3) LAM was more affected by CP. With negative-attribute ratings, CP of LAM was higher for uni- than for bipolar scales. CP was more pronounced for Liked- than Disliked-samples, resulting in higher score fluctuation. CAT was more affected by contrast effects whereas LIN and LAM were more affected by panelist effects. Polarity effects were obvious for the Mild-bitterness sample, showing significantly different results between uni- vs. bipolar scales [3.91 vs. 6.39, 4.28 vs. 6.49, and 41.05 vs. 63.24, respectively, for CAT, LIN and LAM]; all ratings from bipolar scales were not on the negative-side. For the Strong-bitterness sample, unipolar and bipolar ratings were on the negative side, with LAM having more consistent pattern. This study revealed some drawbacks of hedonic scales induced by scale polarity/types and attributes.

Practical Application: This study revealed some advantages and drawbacks of hedonic scales induced by scale polarity/types and attribute questions. It was found that consumers better differentiated negative-attribute ratings. The positive attribute question was more affected by confounding effects. Polarity effects were obvious for a low level of bitterness perception.

Keywords: hedonic scales, negative sensory attributes, unipolar hedonic scales

INTRODUCTION

The traditional 9-point hedonic scale has been used to evaluate like/dislike of products since 1940 but it has shown negative and inaccuracy results for negative sensory attributes such as bitterness. To support a negative attribute classification, Ginane et al. (2011) also included bitterness as a negative attribute tested for a negative hedonic value. “The negative side of hedonic scaling is not as fully differentiated as the positive side” resulting in losing the discriminative power. This may be because consumers are likely to rate opinion about what they like more than dislike (Lawless and Heymann, 1999). This statement was agreed with Caporale et al. (2009) who found that for a highly preferred dish, the more frequently it was served, the less leftover was observed; the contrast result was observed for the least preferred dish such as vegetable. Stein et al. (2003) suggested that rating bitter attribute was highly sensitive for exposure effects. Improper testing due to inappropriate testing protocol, experimental design, questionnaire, target consumer, product types and data analysis could decrease a discriminative power of an experiment.

OBJECTIVES

The aims of this study were to investigating an impact of negative versus positive product attributes with three different scale types (CAT, LAM and LIN) and two polarities (unipolar and bipolar) on degree of liking/disliking. The experiment was classified into 2 sub-objectives as (1) Investigating negative versus positive attribute ratings and (2) Determining polarity effects (unipolar and bipolar scales) for negative attribute ratings.

MATERIALS AND METHOD

Materials and methods: Three commercial grape juices were classified into 3 categories: Welch’s 100% (A), Welch’s light (B), and 50% diluted Welch’s light (C) for testing a positive attribute. Two chicken soups: one with a high level of salt substitution (Potassium Chloride: KCL Mortan® Salt Substitute for salt-free diet, Chicago, IL) at 2% by weight (S) and the other with a mixture of a regular table salt (Sodium Chloride: NaCl Mortan® Salt Substitute, Chicago, IL) and a low level of salt substitution (Potassium Chloride: KCL Mortan® Salt Substitute for salt-free diet, Chicago, IL) at the ratio of NaCl: KCl (2:1) at 1.3 % by weight (M) for testing a negative attribute.

Positive testing: One of the 4 possible random dual serving orders: AB and AC, BA and CA, AC and AB or CA and BA were served. A total of 3 independent sessions derived from three different scale types (9-point categorical scale, 9-point line scale and LAM scale).

Negative testing: Two sets of samples (SM or MS) were served to evaluate a polarity effect (unipolar vs. bipolar).

Procedure: The rating scores were on (1) CAT scale and (2) a LIN scale where 1 refers to dislike extremely and 9 refers to like extremely (Lawless and Heymann, 1998) and (3) a LAM scale (horizontal line) where 1 refers to greatest imaginable dislike and 100 refers greatest imaginable like. The interior phases and space were created according to the published values of Cardello and Schutz (2004). Three attribute questions: overall color (OC), overall taste (OT) and overall liking (OL) were asked for a positive attribute and bitterness perception was asked for a negative attribute.

Statistical analysis: The analysis was carried out with a mixed procedure analysis of variance and once the significant difference ($\alpha < 0.05$) was formed, the follow up analysis was carried out to compare mean pairwisely using multiple comparison test, Tukey’ s Studentized Range (HSD) test (SAS, 2003). The coefficient of determination (R2) (Table1) was calculated from the ANOVA F test table (not shown here) by summing up a total mean square (MS) to use as a denominator [Denominator = ∑MS (Contrast+Treatment+Panel+Residual)]. The R2 of each factor calculated from a fraction of each mean square factor divided by mentioned denominator [R2 of Treatment = MS (Treatment)/ Denominator*100]. The variances of each factor were estimated by Proc Mixed.

RESULTS AND DISCUSSION

(1) The higher the coefficient of determination, the better discriminative power (Hein et al., 2008) and such higher explained variance can be used to justify an important experimental factor. According to Table 4, the treatment effect showed the highest value as expected which implied that the variation of hedonic rating came mostly from product impression rather than biases (contrast, panelist and unexplained variance factors). CAT was ranked first followed by LIN and LAM (90.04 vs. 81.13 and 4.54), respectively; however, a high coefficient value was observed from testing a negative attribute from all three scales. LAM had a very low sensitivity regarding confounding effects (Contrast and Panelist: CP); however, these effects were less pronounced with a negative attribute. Testing disliked samples with CAT was more affected by contrast effects (Table 1); however, the paradox of order was observed. The extremely liked sample tested with LIN was more affected by contrast effects. For moderately liked samples, there was no influence of all three scales. For testing negative attribute, all three scales showed similar performance. Testing a negative attribute has less or no contrast effects involved in either extremely disliked or neither liked nor disliked products but it involved panelist effects (Table 5). However, for positive attribute, CAT and LIN were affected mostly by contrast effects particularly for extremely liked products whereas LAM was prone to panelist effects on both positive and negative attributes. **(2)** For testing negative attribute, LAM was susceptible to CP on both polarities. CAT yielded the highest F value, P value and the lowest in covariance panelist effects which implied a better scale performance (Table 2). Polarity effects were obvious for mild sample (Drewnowski et al., 1997) showing significantly different results between uni- and bipolar scales (Table 3). LAM yielded the consistency pattern of testing negative attributes. The fraction of contrast effects in unipolar was bigger than in the bipolar scale (Table 6). Comparing the polarity effects of negative attributes, CAT was affected by the panelist effect and it was severe when it was tested with a bipolar scale. LIN seemed to have a better score pattern. The CP effect was small which implied a more tolerance to CP effects. LAM was more affected by the panelist effect.

Table 1 Analysis of variance

Positive attribute				
Sample	Contrast	Category	Line	LAM
Good	AB	7.72±0.8	7.68±1.1	78.78±9.8
	BA	7.30±1.2	6.91±1.8	75.03±19.0
	F value	2.71	11.7	1.02
	P value	0.1056	0.0012	0.3179
Moderate	BA	6.53±1.5	6.22±1.7	72.25±19.2
	AB	6.31±1.2	6.30±1.6	77.93±11.7
	F value	0.2	0.02	2.39
	P value	0.6590	0.8858	0.1277
Bad	CA	3.36±1.4	3.70±1.4	31.66±19.2
	AC	3.79±1.6	3.71±1.7	39.67±19.1
	F value	6.24	1.04	2.22
	P value	0.0155	0.3113	0.1422
Negative attribute				
Sample	Contrast	Category	Line	LAM
Mild	MS	6.48±1.9	6.39±1.7	62.92±19.0
	SM	6.29±1.4	6.63±1.6	63.56±19.9
	F value	0.34	0.57	0.22
	P value	0.5619	0.4531	0.6434
Strong	SM	2.67±1.4	2.76±1.7	23.85±15.8
	MS	2.44±1.4	2.78±1.9	24.49±18.2
	F value	0.76	0	0.04
	P value	0.3864	0.9523	0.846

Table 2 Analysis of variance

Unipolar scale				
Sample	Contrast	Category	Line	LAM
Mild	MS	3.87±1.1	4.52±3.9	40.57±10.3
	SM	3.94±0.9	4.06±1.1	41.59±10.4
	F value	0.98	0.79	0.29
	P value	0.3233	0.3757	0.5937
Strong	SM	1.98±0.9	2.06±1.2	22.53±15.0
	MS	1.83±1.0	1.94±1.2	19.95±12.6
	F value	1.11	0	0.99
	P value	0.2942	0.9945	0.3215
Bipolar scale				
Sample	Contrast	Category	Line	LAM
Mild	MS	6.48±1.9	6.39±1.7	62.92±19.0
	SM	6.29±1.4	6.63±1.6	63.56±19.9
	F value	0.34	0.57	0.22
	P value	0.5619	0.4531	0.6434
Strong	SM	2.67±1.4	2.76±1.7	23.85±15.8
	MS	2.44±1.4	2.78±1.9	24.49±18.2
	F value	0.76	0	0.04
	P value	0.3864	0.9523	0.846

Table 3 Analysis of variance

Sample	Scale	Category	Line	LAM
Mild	Bi	6.39±1.7	6.49±1.7	63.24±19.4
	Uni	3.91±1.0	4.28±2.9	41.08±10.3
	F value	176.14	27.02	110.78
	P value	<0.0001	<0.0001	<0.0001
Strong	Bi	2.56±1.4	2.77±1.8	24.17±16.9
	Uni	1.91±1.0	1.99±1.2	21.24±13.8
	F value	16.15	14.19	1.99
	P value	<0.0001	<0.0001	0.1595

Table 4 Coefficient of determination (R2)

Attributes	Scale	R^2			
		T	C	Panel	r
Positive [Bipolar]	Category	90.04*	3.94	1.53	4.49
	Line	81.13	6.49	5.92	6.46
	LAM	4.54	35.32	47.34	12.8
Negative [Bipolar]	Category	99.61	0.03	0.09	0.30
	Line	99.51	0.08	0.01	0.39
	LAM	99.36	0.00	0.23	0.40
Negative [Unipolar]	Category	98.86	0.31	0.38	0.45
	Line	97.29	0.55	0.50	1.66
	LAM	88.60	0.73	10.09	0.58

Table 5 Shared explained variance for each attribute question

Scale	Positive attribute			Negative attribute		
	Sample	R^2		Sample	R^2	
		Contrast	Panelist		Contrast	Panelist
Category	Good	48.86	33.57			
	Moderate	26.39	6.32	Mild	7.03	72.34
	Bad	50.27	4.21	Strong	6.23	85.52
Line	Good	76.29	4.45			
	Moderate	0.46	68.87	Mild	14.18	60.85
	Bad	0.00	23.10	Strong	0.003	2.45
LAM	Good	18.56	63.71			
	Moderate	18.37	74.05	Mild	2.72	4.01
	Bad	51.52	24.28	Strong	2.02	44.75

Table 6 Shared explained variance for each scale polarity

Scale	Unipolar			Bipolar		
	Sample	R^2		Sample	R^2	
		Contrast	Panelist		Contrast	Panelist
Category	Mild	10.20	20.41	Mild	7.03	72.34
Line		36.95	10.39		14.18	60.85
LAM		2.28	89.78		2.72	4.01
Category	Strong	27.70	25.82	Strong	6.23	85.52
Line		20.21	9.55		0.003	2.45
LAM		10.65	78.61		2.02	44.75

CONCLUSION

This experiment concluded that both CAT and LIN yielded similar performance to assess the degree of liking/disliking for either positive or negative attribute; however, LAM can be used for negative attribute. When testing extremely liked product, one should be aware of contrast effects for a positive attribute while less or no contrast effect involved for a negative attribute. Polarity effects were obvious for mild samples. The unipolar scale had higher score fluctuation and was more susceptible to contrast effects. LAM yielded the consistency pattern of testing negative attribute; however, one should be aware of severely panelist effects.

1. Caporale, G., Policastro, S., Tuorila, H. and Monteleone, E. 2009. Hedonic ratings and consumption of school lunch among preschool children. Food Quality and Preference 20(7): 482-489.
2. Cardello, A.V. and Schutz, H.G. 2004. Research Note Numerical scale-point locations for constructing the LAM (Labeled affective magnitude) scale. J. Sensory Studies. 19: 341-346.
3. Drewnowski, A., Henderson, S.A. and Shore, A.B. 1997. Genetic sensitivity to 6-n-proprylthiouracil (PROP) and hedonic responses to bitter and sweet tastes. Chemistry Sense 22: 27-37
4. Ginane, C., Baumont, R. and Favreau-Peigne, A. 2011. Perception and hedonic value of basic tastes in domestic ruminants. Physiol Behav 104(5): 666-674
5. Hein, K.A., Jaeger, S.R., Carr, B.T. and Delahunty, C.M. 2008. Comparison of five common acceptance and preference methods. Food Quality and Preference 19: 651-661.
6. Lawless, H.T., and Heymann, H. 1999. Sensory Evaluation of Food: Principal and Practices. Maryland: Aspen publisher. 827p.
7. Stein, L.J., Nagai, H., Nakagawa, M. and Beauchamp, G.K. 2003. Effects of repeated exposure and health-related information on hedonic evaluation and acceptance of a bitter beverage. Appetite 40(2): 119-129.

Variations in heading ratings characterized by scale polarity scale types and attributes

Jirangrat, W

2012-06-26