Food Quality and Preference 95 (2022) 104377

Contents lists available at ScienceDirect

Food Quality and Preference

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

Short Communication

Sample discrimination through profiling with rate all that apply (RATA) using consumers is similar between home use test (HUT) and central location test (CLT)

Jun Niimi [a,*], Elizabeth S. Collier [b], Lisa-Maria Oberrauter [a], Victoria Sörensen [a], Cecilia Norman [a], Anne Normann [a], Marcus Bendtsen [c], Penny Bergman [a]

[a] Perception & Design Unit, Department of Material and Surface Design, RISE Research Institutes of Sweden, Frans Perssons väg 6, 412-76 Gothenburg, Sweden
[b] Perception & Design Unit, Department of Material and Surface Design, RISE Research Institutes of Sweden, Malvinas väg 3, 114-28 Stockholm, Sweden
[c] Department of Health, Medicine and Caring Sciences, Division of Society and Health, Linköping University, Linköping, Sweden

ARTICLE INFO

Keywords:
RATA
CLT
HUT
Profiling
Consumers

ABSTRACT

The impact of testing locations, home use test (HUT) and central location test (CLT) on consumer evaluations of food products using rate all that apply (RATA) was investigated. As a case study, eight cold cuts were evaluated: four vegetarian/vegan and four meat-based products. A between-subjects design was used, whereby consumers were randomly allocated to either HUT or CLT test location (58 and 71 consumers, respectively). To retain as much similarity as possible across locations, consumers in both groups received identical bags of products with palate cleansers and instructions. Consumers evaluated the products using a lexicon consisting of 32 sensory attributes based on similar studies and benchtop tasting, using RATA with 7pt scales. A total of 30 and 31 attributes differed significantly ($p < 0.05$) across the products for HUT and CLT, respectively. Sample discrimination was similar between the two locations. Location significantly ($p < 0.05$) affected discrimination of 14 attributes, but a particular location having consistently higher attribute means was not observed. Bootstrapping of the attribute means per product showed no significant differences between the two testing locations, and multilevel regression models using Bayesian inference did not reveal marked differences in expected ratings between locations. Further comparisons of sample discrimination patterns through principal component analysis showed that the two locations were very similar, including the overlap of confidence ellipses. The between-subjects design strengthens the results: that comparable sensory profiles were obtained from different consumers in different testing locations supports the notion that RATA data from consumers can be reliably collected for relatively sensorially distinct products with minimal data compromise.

1. Introduction

The COVID-19 global pandemic during the year 2020 has changed how we live as a society. Heightened hygiene and safety concerns, as well as restrictions on gatherings to stop the spread of infection have also affected how we conduct sensory/consumer tests. While typically panellists or consumers may gather in one place during sessions and evaluations, this was temporarily disallowed. However, sensory studies do not always necessarily need to be conducted as a central location test (CLT), instead data can be collected from consumers at home. From a product development perspective, it is arguable that data from home use tests (HUT) have higher ecological validity than that collected from a controlled yet sterile environment, particularly when measuring consumer affect.

HUT is often used to collect hedonic data from consumers. The results obtained from such tests provide valuable feedback for industry. Nowadays, the wide availability of data acquisition software and speed of digitalisation with online tests have made data collection easier than ever. However, comparisons between CLT and HUT have shown contradictory results. While some show no difference in consumer hedonic responses between CLT and HUT (Sinesio, Moneta, Di Marzo, Zoboli, & Abbà, 2021), others reported overall lower scores in hedonics using CLT than HUT with consumers (Wendin, Åström, & Ståhlbröst, 2015). Further, the effect of testing location on consumer responses can depend

* Corresponding author.
E-mail address: Jun.niimi@ri.se (J. Niimi).

https://doi.org/10.1016/j.foodqual.2021.104377
Received 18 August 2021; Received in revised form 30 August 2021; Accepted 31 August 2021
Available online 3 September 2021

on the type of product tested (Boutrolle, Delarue, Arranz, Rogeaux, & Köster, 2007). Clustering patterns within CLT and HUT by liking of products can also vary, shown through a between-subjects design (Zhang, Jo, Lopetcharat, & Drake, 2020). These results are akin to context effects, where consumer responses can also differ between testing in the laboratory, evoked consumption context through imagination whilst being in a laboratory, real life situation (be it a bar, restaurant, as well as at home), and virtual reality (Lichters, Möslein, Sarstedt, & Scharf, 2021; Sinesio et al., 2021).

In most cases, the responses measured nevertheless were hedonics, emotions, and at times willingness to pay. Extending from hedonics, consumers responses to stir fried dishes consumed at either home or in the laboratory has been investigated, where a set of 10 sensory attributes, hedonics, and implicit measures including facial expression, heat rate, and consumption duration were measured (De Wijk et al., 2019). Although only 18 participants were tested, nine of the ten sensory attributes were not affected by test location. Implicit measures, however, were dependent on test location. It is difficult to conclude which context/test location is best to test consumers, as each location has their pros and cons (Wendin et al., 2015), though immersive contexts may yield closer data to actual consumption contexts (Lichters et al., 2021).

The mentioned studies above report alterations in hedonic responses between testing consumers at home and under controlled conditions. It is plausible that other consumer data, e.g., from rapid methods such rate/check all that apply (RATA and CATA), could also be affected by test location. Consumers are increasingly used to profile products using RATA and CATA (Ares, Barreiro, Deliza, Giménez, & Gámbaro, 2010; Ares et al., 2014). These methods have served as a lower resolution alternative to traditional methods such as quantitative descriptive analysis and spectrum method but produce data with relevance to consumers' perception of sensory properties, with reduced timeframes for data acquisition. Although consumers can be tested at home, the impact of test location on profiling results from RATA or CATA has not been extensively investigated. Differences in measurements across test locations, for instance, could be due to the testing locations affecting cognitive processes akin to analytical vs holistic (Prescott, Lee, & Kim, 2011). Consumer responses from CLT may tend to induce more analytical processes whilst HUT may be less so, despite the method. Further, in laboratory-based or central location tests, consumers may be better focused particularly due to sterile environments, whereas unforeseen or uncontrolled distractions at home could affect consumers' concentration. In the case of RATA/CATA, unlike hedonic measurements, variation in profiling responses by test location may be a cause for concern, particularly during product development. The leniency of deviations in sensory properties of products are likely to be held as more stringent by researchers, due to these measures being comparatively more objective than hedonics or emotions. If test location affects profiling data, product developers may have difficulty in deciding actions for refining products based on such data. Knowing the impact of testing location on profiling methods using consumers could be an important asset to researchers and industry alike. If data quality is not compromised, this could present opportunities to continue experimentation even during globally catastrophic events such as the COVID-19 pandemic.

A few recent studies investigated testing locations/contexts on profiling data of CATA and RATA. While data from CATA did not seem to be strongly affected by CLT vs HUT (Lee & Lee, 2021; Schouteten, Gellynck, & Slabbinck, 2021), data from RATA was influenced by context (Sinesio et al., 2021). Both studies used a within-subject design to test across contexts, though learning through multiple exposures may still affect the data. The effect of testing location on data from RATA is thus not yet firmly established. To test the hypothesis that testing context may not impact on profiling data collected by consumers, here the impact of test location on the data quality of RATA on commercially available products was investigated. This study used commercial cold cuts as a case study to determine whether the sensory profile of the products would differ between CLT and HUT, using different groups of consumers for each test location.

2. Materials and methods

2.1. Samples

The products tested were commercial cold cuts available from a local supermarket in Gothenburg, Sweden. A total of eight products were tested, four each of vegetarian and meat-based products; V1 = vegetarian salami with egg protein and hydrolysed vegetable protein, V2 = smoked vegetarian slices and soy protein, V3 = vegan pepperoni made from mycoprotein, V4 = vegan slices made from mycoprotein, M1 = sausage slices, M2 = smoked turkey breast, M3 = pepperoni salami, and M4 = roast beef slices. Samples were selected to represent a range of commercially available cold cuts based on meat and vegan/vegetarian products. Both meat and non-meat products were used as this study was part of a larger project investigating plant-based proteins. Relevant attributes for testing were derived from published studies that have investigated cold cuts and meat substitute products (Aaslyng, Vestergaard, & Koch, 2014; Delahunty, McCord, O'Neill, & Morrissey, 1997; Elzerman, van Boekel, & Luning, 2013; Pham et al., 2008). Benchtop testing was conducted with four researchers to generate further attributes that were relevant but missing from the literature. This process resulted in a total of 32 attributes for RATA: 4 appearance, 9 aroma, 8 texture, and 11 taste/flavour attributes. Definitions were provided for ease of understanding for the consumers along with scale anchors.

2.2. Consumers

A between-subjects design was adopted, where different samples of consumers were recruited per testing location. Due to the restrictions in place controlling for social distancing at the time of testing, testing in booths within a practical timeframe was not possible. Therefore, CLT was conducted in a large conference room to allow for testing multiple people with safe social distancing.

A total of 129 consumers were recruited, and randomly assigned to either the CLT (N = 58; 74.1% female; 25.9% male) or HUT (N = 71; 67.6% females; 32.4% males) test location. These sample sizes are within the range which has been previously shown to have good data stability for RATA (Ares et al., 2014). In both locations, the age ranges of the consumers were evenly distributed between 18 and 65 + years of age. Similar distributions in age groups were seen between the two sample groups; in order of CLT and HUT for each age category, 18–24 years = 6 and 4, 25–34 years = 12 and 8, 35–44 years = 10 and 18, 45–54 years = 13 and 16, 55–64 years = 13 and 14, and 65 + years = 4 and 11, respectively.

2.3. Consumer testing

In both CLT and HUT test locations, the consumers were provided with evaluation bags containing samples, serviettes, cutlery, crackers, attribute list with definitions, and an instruction sheet describing how to access the questionnaire and the tasting protocol. The consumers allocated to CLT performed the test at RISE Research Institutes of Sweden in Gothenburg, in a large room under white light at ambient temperature and tables were spaced 2 m apart from each other. The consumers allocated to HUT collected the package from a set location in Gothenburg, Sweden. HUT consumers were instructed to keep the bags refrigerated at home until they were ready to evaluate the products and evaluate within a week of collection. Samples were evaluated using RATA on 7pt scales. All consumers evaluated the same set of samples by completing questionnaires on smart phones using RedJade (RedJade, Redwood City, USA) and compensated with a gift card.

The study was assessed for compliance with national research ethics standards through an internal process at RISE and was approved by management at the Department of Material and Surface Design. The

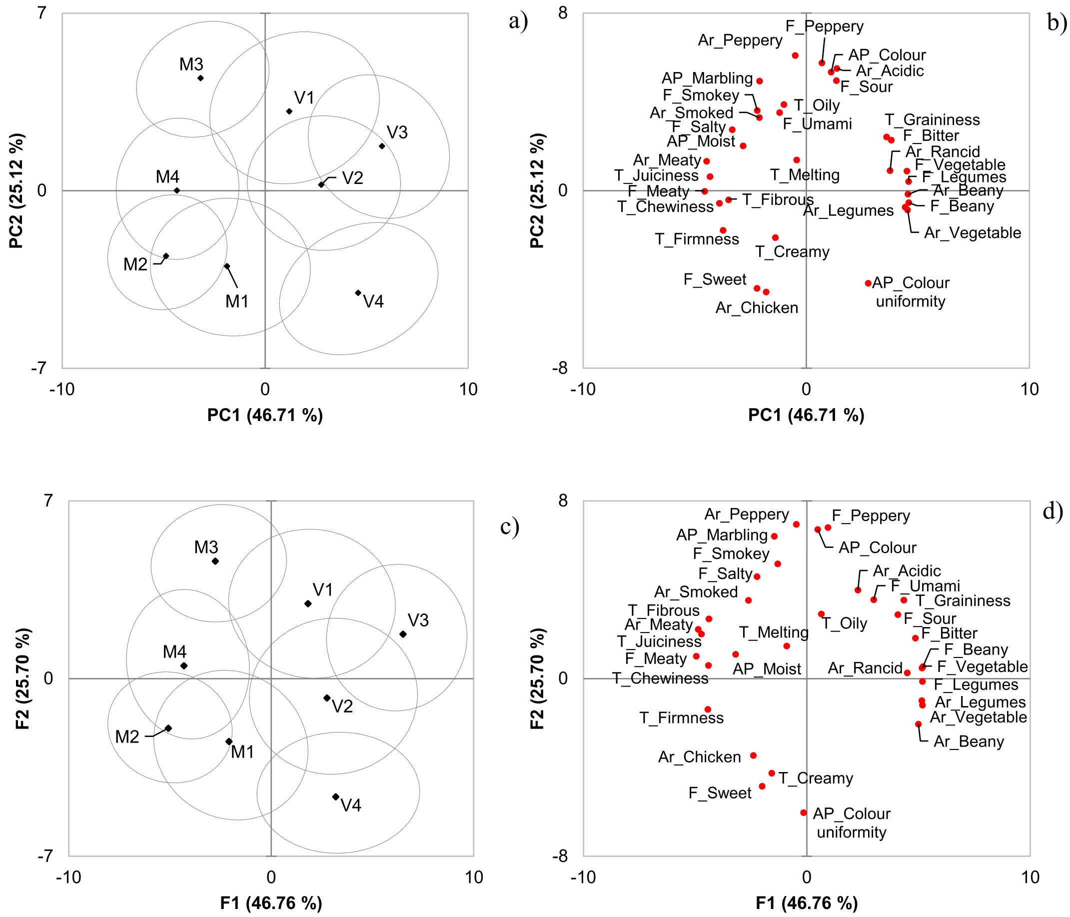

Fig 1. PCA plots of eight cold cuts on the first two PCs assessed at two locations: a) scores CLT, b) loadings CLT, c) scores HUT, and d) loadings HUT.

study was conducted in accordance with the ethical principles described in the Declaration of Helsinki (2013). All participants gave written informed consent to take part and to be recorded before the session started. No invasive methods were used, and no sensitive information was collected. Personal data was collected and processed in accordance with the General Data Protection Regulation (EU) 2016/679 (GDPR).

2.4. Data analysis

Data was pre-processed by replacing unrated attributes with 0 in the data matrices of RATA. The means of attributes for each product were calculated per test location. The mean attribute intensities across all matching products between the CLT and HUT data sets were correlated using Pearson's correlation. Error of the mean attribute intensities from the correlation equation for CLT and HUT data were calculated as Y and X residuals, respectively, and analysed with descriptive statistics. A univariate ANOVA was performed separately on data sets from each location, taking samples and consumers as fixed and random effects, respectively. The F-values associated with attributes when taking samples as a fixed factor were analysed between locations with a student's *t*-test. Sample means of attributes were correlated between the data sets for their correlation coefficients. A further univariate ANOVA was conducted on the entire data set, taking test location and sample as fixed factors, and 2-way interactions thereof. The RATA profiles were bootstrapped (1000 times) and compared each attribute across each sample between the two locations.

To further compare the data sets, principal component analysis (PCA) was performed on the two data sets separately. Scores were plotted with confidence ellipses based on bootstrap resampling of the samples 1000 times. Univariate ANOVA, bootstrapping, PCA, and Peason's correlation were performed using XLstat ver. 2020.05.01 (Addinsoft SARI, Paris, France) at an alpha level of 5%.

Considering the nesting of data within both subjects and products, the data were additionally analysed using multilevel normal regression models to account for these dependences. Models were estimated with Bayesian inference ((Bendtsen, 2018); normal priors for all coefficients: $\mu = 0$; $\sigma = 10$, and exponential priors for errors) using Hamiltonian Monte Carlo ('ulam' function from the 'rethinking' package in R, version 3.6.3, The R Foundation for Statistical Computing Platform).

3. Results

Two groups of consumers evaluated the cold cuts either at home or a central location. The mean intensities of each attribute per sample from CLT data set were correlated against the HUT data (Appendix Fig. A1). The attribute means between the two test locations gave $R^2 = 0.9275$, a fit of $y = 1.179x - 0.3676$, and were significantly correlated ($r(253) = 0.928$, $p < 0.001$). The average residual calculated for CLT and HUT were 0.28 and 0.25pts of the 7pt scale, respectively. Analysis of the data sets individually by location showed 31 and 30 attributes, from a total of 32 attributes, significantly ($p < 0.05$) differed across the products in the CLT and HUT test location, respectively. F-values of attribute discrimination by samples between the two locations matched well for most of the significantly different attributes (Appendix Table A1) and further analysis showed no significant differences in these F-values between the locations ($t(62) = 1.999$, $p = 0.354$). Only four attributes had higher F-values in CLT of more than 10: AP_Colour – $F_{HUT} = 22.2$ and F_{CLT} 61.1; Ar_Peppery – $F_{HUT} = 26$ and F_{CLT} 36.8; T_Chewiness – $F_{HUT} = 32.7$ and F_{CLT} 50.9; F_Peppery – $F_{HUT} = 68.7$ and F_{CLT} 82.5. Three attributes (F_Sour, Sweet, and Umami) had moderate regression correlation

Table 1
Shows the most probable effect (mean of the posterior distribution) comparing CLT and HUT based on multilevel Bayesian inference analysis for each attribute with 95% compatibility intervals. Positive values indicate higher ratings are expected with HUT than CLT, and vice versa for negative values. The posterior probabilities of obtaining any effect, an effect greater than 0.5 pts, and an effect greater than 1 pt in the direction of the mean are also presented.

Category*	Attribute	Mean of posterior distribution (95% CI)	Posterior probability of effect (%)[a]	Posterior probability of effect greater than 0.5pt (%)[a]	Posterior probability of effect greater than 1pt (%)[a]
AP	Moist	0.17 (-0.21 – 0.56)	81	5	< 1
AP	Marbling	0.47 (0.09 – 0.86)	99	45	1
AP	Colour	−0.24 (-0.52 – 0.05)	95	3	< 1
AP	Colour Uniformity	−0.17 (-0.50 – 0.17)	84	2	< 1
Ar	Smoked	0.10 (-0.31 – 0.51)	68	3	< 1
Ar	Beany	0.23 (-0.17 – 0.63)	87	9	< 1
Ar	Peppery	0.10 (-0.32 – 0.53)	67	3	< 1
Ar	Rancid	0.19 (-0.22 – 0.58)	82	6	< 1
Ar	Chicken	0.04 (-0.34 – 0.42)	58	<1	<1
Ar	Vegetable	0.15 (-0.24 – 0.55)	78	4	< 1
Ar	Meaty	−0.25 (-0.63 – 0.15)	90	10	< 1
Ar	Acidic	0.09 (-0.41 – 0.58)	64	5	< 1
Ar	Legumes	0.27 (-0.13 – 0.67)	91	12	< 1
T	Chewiness	0.20 (-0.21 – 0.61)	84	8	< 1
T	Graininess	0.30 (-0.11 – 0.71)	92	17	<1
T	Firmness	0.03 (-0.41 – 0.49)	56	2	< 1
T	Juiciness	−0.53 (-0.93 – −0.15)	99	57	1
T	Melting	−0.74 (-1.21 – −0.28)	99	84	14
T	Fibrous	0.17 (-0.28 – 0.62)	78	8	< 1
T	Creamy	−0.35 (-0.82 – 0.13)	92	26	< 1
T	Oily	0.22 (-0.26 – 0.71)	82	13	< 1
F	Salty	−0.26 (-0.73 – 0.21)	86	16	< 1
F	Peppery	0.03 (-0.36 – 0.43)	56	1	< 1
F	Smokey	−0.03 (-0.51 – 0.45)	55	3	< 1
F	Sour	0.08 (-0.41 – 0.59)	63	5	< 1
F	Sweet	0.22 (-0.24 – 0.68)	83	11	< 1
F	Umami	0.16 (-0.37 – 0.69)	72	10	< 1
F	Beany	0.29 (-0.13 – 0.71)	91	16	< 1
F	Meaty	−0.43 (-0.88 – 0.02)	97	38	< 1
F	Legumes	0.21 (-0.24 – 0.65)	82	10	< 1
F	Bitter	0.12 (-0.34 – 0.56)	69	4	< 1
F	Vegetable	0.26 (-0.24 – 0.76)	85	17	< 1

*AP = appearance, Ar = aroma, T = texture, and F = flavour. [a] In the direction of the mean

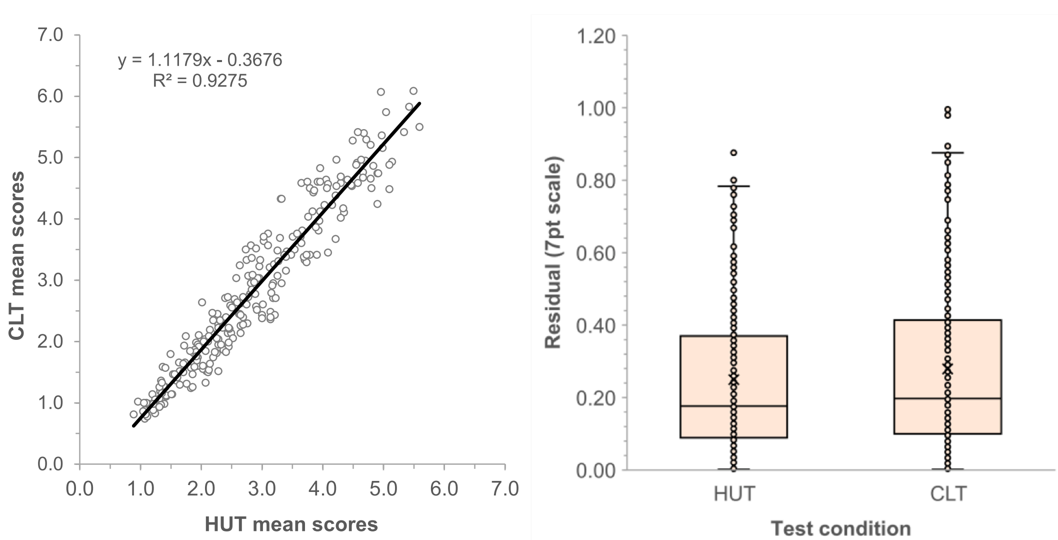

Fig A1. Left: correlation of mean attribute intensities from all samples between CLT and HUT. Right: box and whisker chart of mean attribute residuals for HUT and CLT data (middle "x" denotes for the mean, middle line in the box is the median, upper and lower extremes of the box are upper and lower quartiles, respectively, and upper and lower ends of the whiskers are the maximum and minimum, respectively).

coefficients between locations, while the remainder having a minimum of 0.811. A further 2-way ANOVA showed 30 significantly ($p < 0.01$) different attributes by sample (F_sweet was no longer significant, and F_umami remained non-significant) and 14 significantly ($p < 0.05$) different attributes by location (Appendix Table A2). Most of the differences by locations were within a range of 0.5 on the scale, but with two exceptions for T_Juiciness (0.5) and T_Melting (0.7) (both attributes were rated significantly higher in CLT ($p < 0.001$) and had the highest F-values by location, F = 23.4 and 32.6, respectively). Otherwise, among the 14 significantly different attributes, 7 attributes each were rated

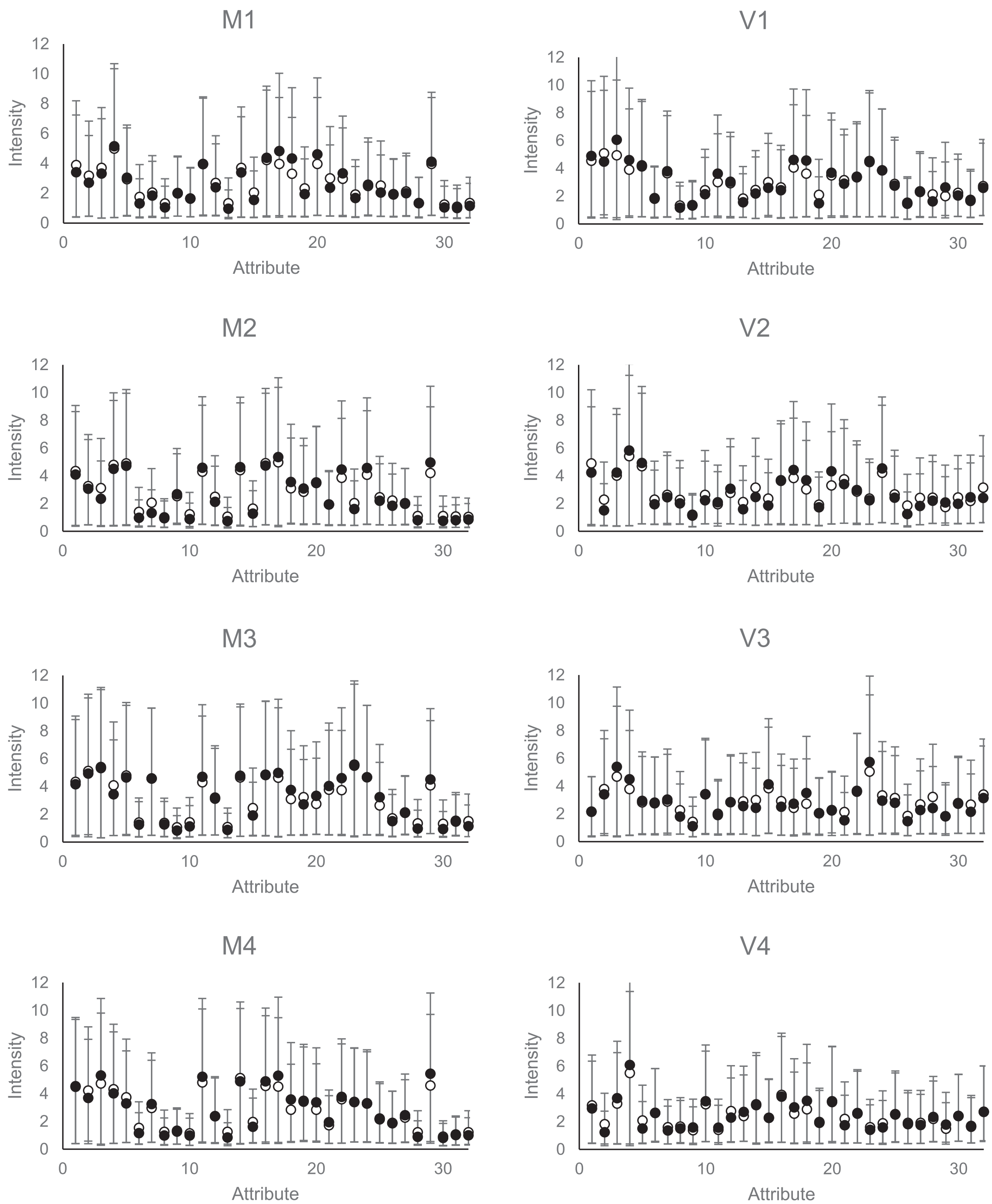

Fig A2. Attribute means for each sample bootstrapped and compared against CLT data (black circles) and HUT (white circles). Bars denote for bootstrapped 95% confidence intervals. M = meat-based product and V = vegetarian/vegan-based product.

significantly higher in CLT and in HUT. A particular location having consistently higher attribute means was not observed. Significant 2-way sample × location interactions were detected for only two attributes (AP_Colour and AP_Colour uniformity), suggesting slight differences in how the samples were scored when assessed in HUT and CLT.

The two data sets were analysed with PCA (Fig. 1). The explained variances on the first PCs for CLT and HUT were 71.8 and 72.5, respectively. The discrimination of scores and loadings projection showed very similar patterns between the locations, namely separation of products by meat and plant-based ingredients. Further, the patterns in the overlaps of confidence ellipses between the samples in both data sets were identical. Slight variations in the projections of a few attributes were seen in how they were discriminated, such as T_Oily and F_Umami along the PC1 and T_Fibrous along PC2.

The data sets were bootstrapped for each sample, separately for the two test locations (Appendix Fig. A2.). The confidence intervals of the resampled attributes per sample overlapped between the ratings measured from CLT and HUT. Bootstrapping suggested therefore that all attributes did not significantly differ between CLT and HUT locations for each sample type, thus no statistically significant differences were observed across testing locations.

For the multilevel regression models, for each attribute, coefficients' posterior mean was calculated and used as a point estimate for the expected difference in ratings (in terms of points on the 7pt scale) moving from CLT to HUT using Bayesian inference as described in section 2.4. The posterior probability of obtaining an effect of testing location (taken as the probability of the difference being less or greater than zero given a negative or positive mean difference respectively), as well as the probability of detecting differences greater than 0.5pts and 1pt in the direction of the mean were also calculated (Table 1). The probability of obtaining effects of testing location was not excessively high, especially considering the small size of the expected differences (posterior means).

Table A1
List of effects of univariate ANOVA run on data sets individually by locaiton with F- and *p*-values, and regression correlation coefficients between ratings at the two locations. *P*-values in **bold** indicate significance at alpha = 5%.

Attribute[§]	HUT				CLT				Correlation coefficients
	Sample		Consumer		Sample		Consumer		
	F	p	F	p	F	p	F	p	
AP_Moist	30.3	**< 0.0001**	5.0	< 0.0001	23.6	**< 0.0001**	4.8	< 0.0001	0.940
AP_Marbling	42.7	**< 0.0001**	3.7	< 0.0001	36.2	**< 0.0001**	4.2	< 0.0001	0.990
AP_Colour	22.2	**< 0.0001**	3.1	< 0.0001	61.1	**< 0.0001**	2.3	< 0.0001	0.928
AP_Colour U*	12.6	**< 0.0001**	3.1	< 0.0001	19.4	**< 0.0001**	2.8	< 0.0001	0.814
Ar_Smoked	25.8	**< 0.0001**	3.9	< 0.0001	27.9	**< 0.0001**	3.6	< 0.0001	0.974
Ar_Beany	9.1	**< 0.0001**	4.7	< 0.0001	13.3	**< 0.0001**	4.7	< 0.0001	0.977
Ar_Peppery	26.0	**< 0.0001**	4.7	< 0.0001	36.8	**< 0.0001**	5.5	< 0.0001	0.975
Ar_Rancid	9.7	**< 0.0001**	7.2	< 0.0001	5.6	**< 0.0001**	5.8	< 0.0001	0.965
Ar_Chicken	11.1	**< 0.0001**	6.9	< 0.0001	12.3	**< 0.0001**	4.9	< 0.0001	0.964
Ar_Vegetable	23.2	**< 0.0001**	4.7	< 0.0001	26.9	**< 0.0001**	3.6	< 0.0001	0.982
Ar_Meaty	37.4	**< 0.0001**	2.4	< 0.0001	40.9	**< 0.0001**	4.3	< 0.0001	0.990
Ar_Acidic	1.6	0.147	6.5	< 0.0001	4.2	**0.0002**	6.3	< 0.0001	0.811
Ar_Legumes	15.6	**< 0.0001**	6.1	< 0.0001	19.5	**< 0.0001**	4.4	< 0.0001	0.950
T_Chewiness	32.7	**< 0.0001**	6.2	< 0.0001	50.9	**< 0.0001**	7.2	< 0.0001	0.969
T_Graininess	13.8	**< 0.0001**	4.6	< 0.0001	19.2	**< 0.0001**	4.5	< 0.0001	0.968
T_Firmness	19.6	**< 0.0001**	5.5	< 0.0001	27.6	**< 0.0001**	5.5	< 0.0001	0.981
T_Juiciness	27.8	**< 0.0001**	6.0	< 0.0001	28.3	**< 0.0001**	3.5	< 0.0001	0.979
T_Melting	2.2	**0.035**	5.5	< 0.0001	3.2	**0.0024**	3.9	< 0.0001	0.930
T_Fibrous	10.7	**< 0.0001**	6.2	< 0.0001	12.0	**< 0.0001**	4.9	< 0.0001	0.898
T_Creamy	6.9	**< 0.0001**	5.0	< 0.0001	9.1	**< 0.0001**	4.5	< 0.0001	0.836
T_Oily	21.3	**< 0.0001**	7.9	< 0.0001	19.8	**< 0.0001**	5.5	< 0.0001	0.920
F_Salty	6.8	**< 0.0001**	7.4	< 0.0001	11.2	**< 0.0001**	5.7	< 0.0001	0.914
F_Peppery	68.7	**< 0.0001**	4.8	< 0.0001	82.5	**< 0.0001**	3.4	< 0.0001	0.990
F_Smokey	19.6	**< 0.0001**	5.4	< 0.0001	23.8	**< 0.0001**	4.0	< 0.0001	0.975
F_Sour	2.8	**0.007**	8.2	< 0.0001	3.9	**0.0004**	6.6	< 0.0001	0.566
F_Sweet	2.1	**0.040**	10.9	< 0.0001	3.2	**0.0024**	11.2	< 0.0001	0.560
F_Umami	1.9	0.060	8.5	< 0.0001	1.8	0.0834	9.2	< 0.0001	0.491
F_Beany	16.8	**< 0.0001**	5.6	< 0.0001	14.4	**< 0.0001**	6.0	< 0.0001	0.925
F_Meaty	39.5	**< 0.0001**	4.2	< 0.0001	44.5	**< 0.0001**	5.0	< 0.0001	0.987
F_Legumes	16.3	**< 0.0001**	6.5	< 0.0001	18.4	**< 0.0001**	5.3	< 0.0001	0.978
F_Bitter	12.1	**< 0.0001**	7.8	< 0.0001	11.3	**< 0.0001**	6.5	< 0.0001	0.925
F_Vegetable	25.1	**< 0.0001**	7.3	< 0.0001	24.0	**< 0.0001**	5.7	< 0.0001	0.971

*AP_Colour uniformity
[§] AP = appearance, Ar = aroma, T = texture, and F = flavour.

Attributes with high probability of any effect (e.g., AP_Marbling, AP_Colour, T_Juiciness, T_Melting) did not represent marked expected deviations across groups, given the small differences expected even at the extremes of the confidence – or compatibility – intervals (95% CI). Finally, the probability of obtaining an effect of more than 1pt difference was very low for most attributes. This means that it is reasonable to consider the data obtained in the HUT and CLT locations similar.

4. Discussion

RATA has been gaining popularity in consumer research, particularly in food and beverages. Results from these methods have been shown to be comparable with descriptive analysis (Danner et al., 2018; Nishida, Lestringant, Cantu, & Heymann, 2021; Oppermann, de Graaf, Scholten, Stieger, & Piqueras-Fiszman, 2017), but is nevertheless a comparatively lower resolution method. The residual error in the data sets observed was inevitable for several obvious reasons; measurements made on humans will have variation, the study was conducted as a between-subjects design, and the subjects were naïve consumers.

The results of the current study showed that RATA profiles are stable across testing locations at the consumer group level. Although hedonics can be affected by testing location and context (Bastian, Danner, Niimi, Ristic, & Johnson, 2019; Stelick & Dando, 2018), consumers could adequately profile the sensory characteristics of cold cuts regardless of test location. The results therefore concur with literature reporting that rapid methods can be used at home (Lee & Lee, 2021; Schouteten et al., 2021) and extends to RATA by consumers. This provides important evidence that future testing of products at home using RATA is a viable option, at least for those requiring little to no preparation by consumers and when including samples with large sensory differences. The consistency of the data should however be confirmed with other products to determine the robustness of the findings of the current study and product dependence on data quality between CLT and HUT.

One limitation of the current study was the lack of comparison to data collected in laboratory booths. Better discrimination could be expected from evaluations in booths due to conditions that may encourage better focus and concentration. The time of day at which a consumer evaluated a product and potential distractions were not controlled, and perceptions may have differed throughout the course of a day, adding noise to the data. Compared to CLT, control over the state of the sample was not possible with HUT. This lack of control in sample serving, such as temperature, may be an additional source of noise in the data. It may be seen that the design of the experiment using between-subjects was a limitation, due to the variation of consumers. Attributes may have been understood differently across consumer groups and lower rates of selection for an attribute itself being low may have skewed the calculation of their means. This may also explain how three attributes were projected slightly differently on the PCAs. Nonetheless, despite the design of the experiment, the strong correlations detected, together with the other analyses, these results suggest that the data from RATA can be collected from consumers at home with little compromised data quality. Furthermore, a benefit of the between-subjects design was the avoidance of any bias that could have occurred from a within-subject design, such as learning effects or demand characteristics from repeated evaluations.

It is common practice to collect consumer affect data through HUT, but the same – and additional – considerations still apply to profiling. As profiling is reliant on the consumers ability to discriminate sensory attributes, the products tested need to suit the assumed level of proficiency

Table A2
List of effects of 2-way ANOVA with interactions run on the data sets combined and analysed with sample and location as effects. *P*-values in **bold** indicate significance at alpha = 5%.

Attribute[§]	Sample		Location		Sample × Location	
	F	*p*	F	*p*	F	*p*
AP_Moist	34.8	**< 0.0001**	2.8	0.0967	1.1	0.3732
AP_Marbling	57.4	**< 0.0001**	16.2	**< 0.0001**	0.4	0.9174
AP_Colour	59.4	**< 0.0001**	6.5	0.0109	5.2	**< 0.0001**
AP_Colour U*	23.1	**< 0.0001**	2.4	0.1224	2.8	**0.0067**
Ar_Smoked	39.1	**< 0.0001**	0.6	0.4381	0.6	0.7177
Ar_Beany	14.8	**< 0.0001**	4.2	**0.0415**	0.3	0.9430
Ar_Peppery	40.4	**< 0.0001**	0.7	0.3902	0.9	0.4967
Ar_Rancid	8.5	**< 0.0001**	3.2	0.0737	0.2	0.9735
Ar_Chicken	14.2	**< 0.0001**	0.1	0.7515	0.4	0.9041
Ar_Vegetable	35.0	**< 0.0001**	1.7	0.1936	0.5	0.8701
Ar_Meaty	60.4	**< 0.0001**	4.0	**0.0468**	0.4	0.8829
Ar_Acidic	3.0	**0.0042**	0.4	0.5287	0.5	0.8087
Ar_Legumes	21.8	**< 0.0001**	5.7	**0.0170**	0.7	0.7042
T_Chewiness	46.3	**< 0.0001**	3.7	0.0555	1.1	0.3405
T_Graininess	22.3	**< 0.0001**	6.2	**0.0132**	0.7	0.6331
T_Firmness	28.9	**< 0.0001**	0.1	0.7755	0.5	0.8654
T_Juiciness	37.0	**< 0.0001**	23.4	**< 0.0001**	0.4	0.8721
T_Melting	3.6	**0.0008**	32.6	**< 0.0001**	0.2	0.9779
T_Fibrous	13.6	**< 0.0001**	2.0	0.1577	0.8	0.5805
T_Creamy	10.1	**< 0.0001**	6.8	**0.0092**	1.1	0.3611
T_Oily	22.9	**< 0.0001**	3.2	0.0757	1.0	0.4311
F_Salty	10.2	**< 0.0001**	4.7	**0.0306**	0.8	0.5701
F_Peppery	106.2	**< 0.0001**	0.1	0.7395	1.1	0.3835
F_Smokey	28.8	**< 0.0001**	0.1	0.8134	0.6	0.7809
F_Sour	2.9	**0.0049**	0.5	0.4835	0.9	0.5373
F_Sweet	1.8	0.0826	4.9	**0.0275**	0.5	0.8137
F_Umami	1.4	0.1994	1.3	0.2602	0.5	0.8500
F_Beany	18.4	**< 0.0001**	6.6	**0.0102**	0.7	0.6299
F_Meaty	57.3	**< 0.0001**	10.9	**0.0010**	0.6	0.7728
F_Legumes	20.6	**< 0.0001**	3.3	0.0693	0.2	0.9723
F_Bitter	12.5	**< 0.0001**	1.0	0.3198	0.5	0.8443
F_Vegetable	27.8	**< 0.0001**	4.3	**0.0394**	0.4	0.8979

*AP_Colour Uniformity
[§] AP = appearance, Ar = aroma, T = texture, and F = flavour.

of the judges, that being lower than a trained panel. The appropriateness of testing product prototypes with subtle differences should therefore be carefully considered. Sample discrimination may still be difficult for consumers even though discrimination of sensory attributes of model double emulsions with subtle differences using RATA were comparable to a descriptive analysis panel (Oppermann et al., 2017). Products that require more preparation (e.g., products that need cooking, or that require serving at specific temperatures) also need careful consideration using HUT RATA, as this would undoubtedly introduce noise to the data. Of course, with evaluation bags of products prepared in the current study, serving regime will be affected and monadic serving cannot be ensured. Although the study did not test the effect of providing attribute definitions to the consumers, provision of instructions to consumers is beneficial (Moskowitz, 1996). The caveat to this is that the task at hand may become overwhelming to the consumers if too much information is provided. The research field would benefit from future studies that investigate the optimal amount of information for each measurement method involving consumers when data is collected at their homes.

5. Conclusion

For the testing of refrigerated cold cuts, the results of sensory profiling with RATA using consumers were very similar across CLT and HUT. Test location did not strongly or consistently affect sample discrimination. This signifies that for products where no preparation is involved by the consumers, the profiling results from a central location and at home were comparable. The appropriateness of testing RATA in a HUT setting, however, is likely to highly depend on the choice of product for testing and the amount of preparation required for the evaluation.

CRediT authorship contribution statement

Jun Niimi: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing, Project administration. **Elizabeth S. Collier:** Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing, Analysis, Project administration. **Lisa-Maria Oberrauter:** Investigation. **Victoria Sörensen:** Investigation. **Cecilia Norman:** Investigation. **Anne Normann:** Investigation, Conceptualization. **Marcus Bendtsen:** Investigation, Formal analysis, Writing – review & editing. **Penny Bergman:** Conceptualization, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

References

Aaslyng, M. D., Vestergaard, C., & Koch, A. G. (2014). The effect of salt reduction on sensory quality and microbial growth in hotdog sausages, bacon, ham and salami. *Meat Science, 96*(1), 47–55.

Ares, G., Barreiro, C., Deliza, R., Giménez, A., & Gámbaro, A. (2010). Application of a check-all-that-apply question to the development of chocolate milk desserts. *Journal of Sensory Studies, 25*(s1), 67–86.

Ares, G., Bruzzone, F., Vidal, L., Cadena, R. S., Giménez, A., Pineau, B., et al. (2014). Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA). *Food Quality and Preference, 36*, 87–95.

Bastian, S. E. P., Danner, L., Niimi, J., Ristic, R., & Johnson, T. E. (2019). Alcoholic beverages in context. In H. L. Meiselman (Ed.), *Context*. Cambridge, MA: Woodhead Publishing.

Bendtsen, M. (2018). A gentle introduction to the comparison between null hypothesis testing and bayesian analysis: Reanalysis of two randomized controlled trials. *Journal of Medical Internet Research, 20*(10), e10873. https://doi.org/10.2196/10873

Boutrolle, I., Delarue, J., Arranz, D., Rogeaux, M., & Köster, E. P. (2007). Central location test vs. home use test: Contrasting results depending on product type. *Food Quality and Preference, 18*(3), 490–499.

Danner, L., Crump, A. M., Croker, A., Gambetta, J. M., Johnson, T. E., & Bastian, S. E. P. (2018). Comparison of Rate-All-That-Apply and Descriptive Analysis for the Sensory Profiling of Wine. *American Journal of Enology and Viticulture, 69*(1), 12–21.

De Wijk, R. A., Kaneko, D., Dijksterhuis, G. B., van Zoggel, M., Schiona, I., Visalli, M., et al. (2019). Food perception and emotion measured over time in-lab and in-home. *Food Quality and Preference, 75*, 170–178.

Delahunty, C. M., McCord, A., O'Neill, E. E., & Morrissey, P.a. (1997). Sensory characterisation of cooked hams by untrained consumers using free-choice profiling. *Food Quality and Preference, 8*(5–6), 381–388.

Elzerman, Johanna E., van Boekel, Martinus A. J. S., & Luning, Pieternel A. (2013). Exploring meat substitutes: Consumer experiences and contextual factors. *British Food Journal, 115*(5), 700–710.

Lee, N., & Lee, J. (2021). Comparison of Home Use Tests with Differing Time and Order Controls. *Foods, 10*(6), 1275.

Lichters, Marcel, Möslein, Robert, Sarstedt, Marko, & Scharf, Andreas (2021). Segmenting consumers based on sensory acceptance tests in sensory labs, immersive environments, and natural consumption settings. *Food Quality and Preference, 89*, 104138. https://doi.org/10.1016/j.foodqual.2020.104138

Moskowitz, Howard R. (1996). Experts versus consumers: A comparison. *Journal of Sensory Studies, 11*(1), 19–37.

Nishida, M., Lestringant, P., Cantu, A., & Heymann, H. (2021). Comparing classical descriptive analysis with modified descriptive analysis, modified rate-all-that-apply, and modified check-all-that-apply. *Journal of Sensory Studies, n/a(n/a)*, Article e12684.

Oppermann, A. K. L., de Graaf, C., Scholten, E., Stieger, M., & Piqueras-Fiszman, B. (2017). Comparison of rate-all-that-apply (RATA) and descriptive sensory analysis (DA) of model double emulsions with subtle perceptual differences. *Food Quality and Preference, 56*, 55–68.

Pham, A. J., Schilling, M. W., Mikel, W. B., Williams, J. B., Martin, J. M., & Coggins, P. C. (2008). Relationships between sensory descriptors, consumer acceptability and volatile flavor compounds of American dry-cured ham. *Meat science, 80*(3), 728–737.

Prescott, John, Lee, Soh Min, & Kim, Kwang-Ok (2011). Analytic approaches to evaluation modify hedonic responses. *Food Quality and Preference, 22*(4), 391–393.

Schouteten, J. J., Gellynck, X., & Slabbinck, H. (2021). Do fair trade labels bias consumers' perceptions of food products? A comparison between a central location test and home-use test. *Sustainability, 13*(3), 1384.

Sinesio, Fiorella, Moneta, Elisabetta, Di Marzo, Sabrina, Zoboli, Gian Paolo, & Abbà, Silvia (2021). Influence of wine traits and context on liking, intention to consume, wine-evoked emotions and perceived sensory sensations. *Food Quality and Preference, 93*, 104268. https://doi.org/10.1016/j.foodqual.2021.104268

Stelick, Alina, & Dando, Robin (2018). Thinking outside the booth—the eating environment, context and ecological validity in sensory and consumer research. *Current Opinion in Food Science, 21*, 26–31.

Wendin, K., Åström, A., & Ståhlbröst, A. (2015). Exploring differences between central located test and home use test in a living lab context. *International Journal of Consumer Studies, 39*(3), 230–238.

Zhang, M. T., Jo, Y., Lopetcharat, K., & Drake, M. A. (2020). Comparison of a central location test versus a home usage test for consumer perception of ready-to-mix protein beverages. *Journal of Dairy Science, 103*(4), 3107–3124.