Applied Food Research 5 (2025) 101494

Contents lists available at ScienceDirect

Applied Food Research

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

Exploring sensory profiling with rate-all-that-apply in comparison to check-all-that-apply with and without just-about-right in product development

Vera van Stokkom [a,*], Safira Benguedda [a], Elise Pelgrim [b], Marieke Nijmeijer [a]

[a] Inholland University of Applied Sciences, Agri, Food & Life Sciences, Delft, The Netherlands
[b] Wageningen University, Sensory Science and Eating Behaviour, Wageningen, The Netherlands

ARTICLE INFO

Keywords:
Sensory Research
Rapid Profiling Methods
Just-About-Right
Check-All-That-Apply
Rate-All-That-Apply
Product Development

ABSTRACT

Combining rapid sensory profiling methods Check-All-That-Apply (CATA) or Rate-All-That-Apply (RATA) with Just-About-Right (JAR) might provide insights on attribute presence and intensity, which can be useful during product development. We explored combining RATA and JAR in comparison to CATA, CATA-JAR and RATA. Participants (n=73) were randomly divided into four conditions: CATA (n=18), RATA (n=18), CATA-JAR (n=19) and RATA-JAR (n=18). Participants received three different bean snacks in random order. Each condition included 24 product attributes and questions about the length and difficulty of the questionnaire. Attribute selection frequency differed significantly for all three products between conditions RATA-JAR (71.3 %-78.5 %), RATA (74.3 %-78.2 %) and CATA (52.3 %-54.9 %), CATA-JAR (54.5 %-60.1 %). The percentage of attributes with a significant difference between samples, differed slightly between conditions (RATA-JAR: 29.2 % versus CATA: 29.2 %, CATA-JAR: 37.5 %, RATA: 33.3 %). Correspondence analysis based on the selection frequency contingency tables showed large similarities between conditions. The rating-task was not correlated to the JAR-task for all attributes. In RATA-JAR significantly more attributes were selected as 'too little' compared to CATA-JAR. JAR can provide additional input for product optimization when combined with CATA or RATA. RATA-JAR adds value when information on intensity and the liking of that intensity is needed. However, the added benefit of RATA-JAR compared to CATA-JAR is questionable and might even distort the input for product development, as the selection of the majority of attributes might make it more difficult to pinpoint which attributes are essential. Future research with more participants and products from different categories is recommended.

1. Introduction

Several rapid sensory methods have been developed that are generally faster and provide more flexibility than traditional sensory methods and might therefore be more suitable to use in the product development process, where often fast feedback on products is desired. These rapid methods do not require (a lot of) training and monitoring and therefore an untrained consumer panel can be used (Ares & Jaeger, 2015; Delarue et al., 2014; Delarue, 2015; Jaeger & Ares, 2015; Varela & Ares, 2012). The use of an untrained panel is an advantage for small and medium sized enterprises (SMEs) as resources such as access to expertise, time and equipment are often limited (Howieson et al., 2014). This might be a reason why SMEs apply sensory evaluation methods during product development less frequently compared to larger companies (Świąder & Marczewska, 2021).

Before conducting the current study, a preliminary study was conducted (Pelgrim, 2022). Structured interviews were held with 10 product developers from SMEs based in the Netherlands (38 SMEs were approached). The aims of this preliminary study were to gain understanding of current sensory research practices in SMEs, identify sensory research needs and wants and objectives, and to evaluate which descriptive sensory methods theoretically fulfil these needs. After informed consent was obtained, the semi-structured interview protocol consisted of first talking about the company's new product development (NPD) process by comparing it to the Cooper's stage model (Cooper, 2010). Then, open questions were asked regarding sensory research within the company, divided in the following topics: sensory methods, sensory panel, sensory data handling, challenges in sensory research,

* Corresponding author.
E-mail address: Vera.vanStokkom@inholland.nl (V. van Stokkom).

https://doi.org/10.1016/j.afres.2025.101494
Received 7 August 2025; Received in revised form 27 October 2025; Accepted 4 November 2025
Available online 7 November 2025

improvements in sensory research, failed products, and value of sensory research. Interviews were transcribed and analysed via content analysis – deductive initial coding based on the pre-determined topics followed by line-by-line coding. Most interviewees were familiar with the more traditional sensory methods such as triangle tests and quantitative descriptive analysis and regularly used discrimination tests, preference tests or hedonic scales. The interviewees were unfamiliar with rapid sensory methods, often did not follow basic sensory guidelines such as blinding, randomizing and coding, and had limited knowledge on statistical analyses of sensory data beyond calculating means. Needs and want statements were mostly related to available resources: methods should be suitable for small panel sizes (as small as possible), easy and fast in organising, suitable for multiple objectives and have an easy analysis (as easy as calculating a mean). Based on the findings a combination of Check-All-That-Apply (CATA) and Just-About-Right (JAR) or Rate-All-That-Apply (RATA) and JAR seemed to suit most needs and wants of the interviewees and were considered worth investigating further.

CATA has been investigated, compared to other methods and applied in numerous studies and is proven to be simple and reliable, however, it is limited in its ability to discriminate between samples (Ares et al., 2010; Ares & Jaeger, 2013; Ares, Tárrega et al., 2014; Ares & Jaeger, 2015; Bruzzone et al., 2011; Dooley et al., 2010; Jaeger et al., 2023; Kim et al., 2023; Wang et al., 2023). RATA differs slightly from CATA by also providing intensity ratings of product characteristics, using 3- or 5 point intensity categories. Some research shows that RATA is more comparable to descriptive analysis than CATA and/or has better discrimination between products (Ares, Bruzzone et al., 2014; Kim et al., 2023; Pineau et al., 2022), but not all studies support this (Vidal et al., 2018). A limitation of both CATA and RATA in product development is the lack of information on (the acceptance of) attribute intensities.

In the preliminary study (Pelgrim, 2022), CATA-JAR and RATA-JAR theoretically seemed suitable to use during product development. The methods are still rapid and easy to execute, but JAR adds insights on consumers' acceptance towards the perceived attributes, which can help product development and optimization. The CATA-JAR combination has been studied by Lee et al. (2021) and Jaeger et al. (2015), and both studies showed positive effects on discriminative ability compared to CATA. RATA-JAR might also be a flexible solution giving even more insights for product development as it includes both ratings (RATA) and how these ratings are perceived (JAR), which are valuable insights for product developers. In this study we explored differences in product sensory insights obtained from RATA-JAR, CATA-JAR, CATA and RATA in a consumer test with three bean snacks to better understand if RATA-JAR could be of added value for product development in comparison to CATA, RATA and CATA-JAR.

2. Methods

2.1. Products

For this study we contacted a startup company which was not included in the preliminary interviews, GemBites, a producer of bean snacks made out of field beans. They provided commercially available bean snacks, that were in a taste optimization process. GemBites was not involved in any other way. Three bean snacks were used for this study; two snacks made of oven-dried field beans with different mixtures of herbs from GemBites (GemBites, 2024), GemBites Paprika-Garlic (GBPG) and GemBites BBQ (GBBBQ). The other snack was a more established commercially available snack made of chickpeas smoked paprika flavour from Bites We Love (BWLSP). The two GemBites products were prepared and sealed one day before sensory testing and stored at room temperature. Bite We Love was purchased a few days before sensory testing. All product packages were opened on the sensory test day. The participants got six pieces of every product, served in small disposable plastic containers. The products were given random three-digit codes and in random order.

2.2. Attribute selection

The attributes were determined based on a single one hour session with nine consumers and a discussion among six sensory experts about the attributes generated in the consumer session. All consumers were students at non-food related studies and did not participate in the next consumer sensory test. This resulted in 24 attributes categorised in 'appearance', 'taste and smell' and 'texture': Appearance: puffiness, uniform, visible herbs; Taste and smell: (high) odour intensity, (high) flavour intensity, paprika, sour, garlic, bean-like, bitter, sweet, salty, spicy, nutty, BBQ, savoury, mealy; Texture: Rough, loose peels, crunchy, hard, tough, dry, fatty.

2.3. Participants

Lee et al. (2021) showed that the number of panellists required for a stable result was lower in CATA-JAR (25 consumers) and RATA (30 consumers) compared to CATA alone (50 consumers). Due to the exploratory nature of the study and due to practical reasons we aimed to include a 120 participants, divided over four conditions, i.e. 30 participants per condition. A screener was sent to 2031 consumers in a database from a consumer research agency (Essensor, the Netherlands). The screener was completed by 557 consumers. Consumers were eligible to participate if they met the following criteria: not working in sectors: food, market research, marketing or journalism. Additionally, there had to be at least four weeks between the last participation in research on savoury snacks and they had to eat savoury snacks at least once a month. In total, 403 consumers met the inclusion criteria and received study information on duration, general procedures, tasks, voluntary nature and data handling before giving informed consent by selecting 'I agree and would like to participate in this study'. In total 126 consumers participated, 53 at the test location in Vlaardingen and 73 at the test location in Utrecht. Consumers in the database participate in sensory tests once every two weeks on average and are familiar with general sensory test procedures.

Participants received a 10 euro gift voucher via de consumer research agency for their participation. Unfortunately, GemBites had a production error which resulted in an altered texture on the test day in Vlaardingen, therefore, the results from this test day were excluded. Further description and analysis are based on results from test day in Utrecht (n=73). Of the 73 participants included in the analyses, 50,7 % (n=37) identifies themselves as male and 49,3 % (n=36) as female. All participants were between 23 and 65 years old with a mean age of 48 years (± 13.1). All participants eat savoury snacks at least once a month, whereas 86.3 % consumes savoury snacks daily. Participants were randomly assigned to one of four conditions: CATA n= 18, CATA-JAR n=19, RATA n=18, RATA-JAR n=18. The study falls under an anonymous sensory testing with commercially available products – single sensory session umbrella protocol, which was approved by the ethics committee of Inholland (Umbrella Protocol Sensory and Consumer Research RIC AFL 2021).

2.4. Conditions & procedure

Four different conditions were compared: CATA (n=18), CATA-JAR (n=19), RATA (n=18) and RATA-JAR (n=18). In the CATA condition, participants needed to express if each attribute was present or not present. In the RATA condition, the option 'present' was divided into 'low', 'medium' and 'high'. In the CATA-JAR condition a JAR question appeared for each attribute that was selected as present with three categories: 'too little', 'just about right' and 'too much'. In the RATA-JAR condition, if an attribute was selected as 'low', 'medium' or 'high', JAR appeared to rate the intensity of the attribute with the same three options as with CATA-JAR (Table 1). All product-related questions

Table 1
Overview of the tasks in the four conditions.

Condition	CATA	CATA-JAR	RATA	RATA-JAR
CHECKING Present - not present				
RATING Low-medium-high Only for checked attributes				
JAR Too little – just about right – too much Only for checked attributes				

of all conditions were forced choice. To get an idea of the acceptability of participants on the RATA-JAR questionnaire, a question on length (5 point from very short to very long) and simplicity (5 point from very easy to very difficult) was added. A between-subjects design was used.

The sensory research was carried out at the sensory research facility of Essensor in sensory booths (Central Location Test- CLT). Participants were randomly divided over the four conditions. Samples were handed out simultaneously when all participants in all conditions were ready for the next sample. When finished, participants could leave the room and go home. All the questionnaires were created in EyeQuestion (version 15). Data was collected anonymously.

2.5. Data analysis

The resulting data were statistically analysed with IBM software SPSS Statistics version 30. A significance level of 0.05 was used for all analyses. Frequency tables were made for all conditions for all products and attributes. Chi-Square by transforming data from conditions RATA and RATA-JAR into dichotomous data, i.e. when participants selected either 'low', 'medium' or 'high', the attribute was indicated as present. Total selection frequency percentage were compared between conditions per product per attribute with Chi-Square Test and Fisher's Exact Test. To analyse differences of attribute selection frequency between products within a condition, Chi-Square and Fisher's Exact test were used for CATA and CATA-JAR and Friedman and Wilcoxon tests were used for RATA and RATA-JAR. Total percentage of attributes showing significant differences between products was calculated per condition, using the ordinal data for RATA and RATA-JAR. Correspondence analysis (CA) based on the selection frequency contingency tables was applied to visualize the relationship between the samples and attributes for each condition. Attributes that were selected 50 % or less in all conditions for all products were removed for the CA (only sour). CA was based on binary data for CATA and CATA-JAR and on ordinal data for RATA and RATA-JAR. EyeOpenR® version 6.1.2.6 was applied for visualisation of the CA-biplots. Kruskal Wallis and Mann-Whitney U tests were used to compare the JAR results between CATA-JAR and RATA-JAR and the distribution RATA rating scores between RATA and RATA-JAR. Spearman's rank-order correlation was used to assess correlations between RATA and JAR from the RATA-JAR condition. Length and difficulty between conditions was analysed with Chi-Square.

3. Results

3.1. Attribute selection frequency

Attribute selection frequency increased significantly between the CATA/CATA-JAR (GBPG: 54.9 %, 59.9 %, GBBBQ: 49.3 %, 54.4 %, BWLSP: 52,3 %, 60.1 %) and RATA/RATA-JAR (GBPG: 78.2 %, 78.5 %, GBBBQ: 74.3 %, 71.3 %, BWLSP: 74.5 %, 75.0 %), but not between CATA and CATA-JAR or RATA and RATA-JAR within each product (Table 2).

The effect of the conditions on selection frequency was similar for the three products as visualised in spider plots (Fig. 1). CATA(-JAR) and RATA(-JAR) differed significantly in attribute selection frequency for most attributes (Appendix A). There were no significant differences in selection frequency between conditions for bean-like, nutty, hard and crunchy. CATA and CATA-JAR only differed in selection frequency of sweet for product GBPG. Between RATA and RATA-JAR, four attributes in one product (GBBBQ) showed a significant difference: BBQ, savoury, tough and dry. Total percentage of attributes that show a significant difference in selection frequency between products differed between conditions: CATA 29.2 %, CATA-JAR 37.5 %, RATA 33.3 %, RATA-JAR 29.2 %. When RATA and RATA-JAR were analysed as binary data, this percentage was lower, RATA: 25 % and RATA-JAR 12 % (data not shown). In the CA (Fig. 2) regarding the sensory profiles obtained from the different conditions the biplot explains 100 % of the variance due to there being only three products. Dimension 1 explained a large part of the variance (CATA: 73.1 %, CATA-JAR: 79.8 %, RATA: 91.9 %, RATA-JAR: 85.8 %). In all conditions, GBPG and GBBBQ are clearly different from BWLSP.

3.2. Rating and JAR

There were no significant differences in rating scores between RATA and RATA-JAR for any of the products (Table 3).

Selection of 'Just about right' and 'too much' did not differ significantly between CATA-JAR and RATA-JAR. 'Too little' was selected significantly more often in RATA-JAR than in CATA-JAR for all three products (+20.1 % for GBPG, +12.2 % for GBBBQ, and +17.4 % for BWLSP), which is probably a result of selecting 'low' intensity attributes in CATA less frequently (Table 4, Fig. 3).

RATA and JAR were significantly correlated for the majority of attributes (66.7 % of the attributes for GBPG, 70.8 % for GBBBQ and 66.7 % for BWLSP), but not always. This means that a rating of 'low' was not always correlated to a rating of 'too little', 'medium' not always with 'just about right', and 'high' not always with 'too much' (Appendix B).

Table 2
The average percentage of total attributes selected per product per condition.

	CATA (n=18)	CATA-JAR (n=19)	RATA (n=18)	RATA-JAR (n=18)
GBPG	54.9 %[a]	59.9 %[a]	78.2 %[b]	78.5 %[b]
GBBBQ	49.3 %[a]	54.4 %[a]	74.3 %[b]	71.3 %[b]
BWLSP	52.3 %[a]	60.1 %[a]	74.5 %[b]	75.0 %[b]

Different letters, [a-b] indicate significant differences between conditions within a product.

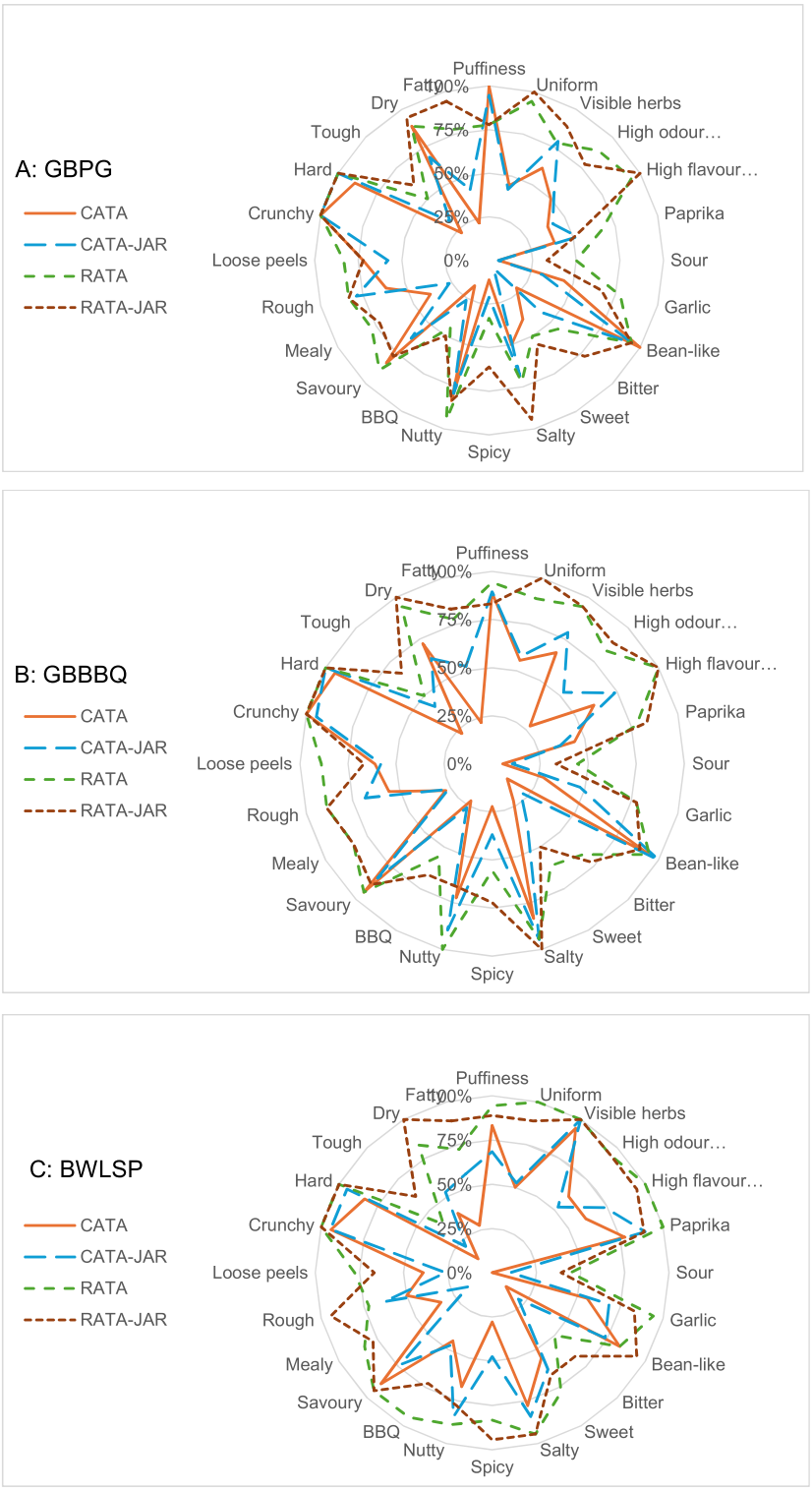

Figure 1. Spider plots with the percentages of attribute selection frequency for conditions CATA, CATA-JAR, RATA and RATA-JAR for three products: GBPG (A), GBBBQ (B), and BWLSP (C).

3.3. Length and simplicity of conditions

During the test it was noticeable for the researchers that participants in the RATA-JAR condition took a bit longer to assess products, however there were no significant differences between the perceived length of the conditions by the participants. The length was mostly rated as 'good' for all conditions: CATA (88.9 %), CATA-JAR (84.2 %), RATA (100 %) and RATA-JAR (83.3 %). RATA-JAR simplicity significantly differs from the other conditions; RATA-JAR was significantly less selected as 'very easy' (16.7 %) and significantly more selected as 'a bit difficult' (22.2 %). The most selected option for RATA-JAR was 'a bit easy' (33.3 %). CATA (50 %), CATA-JAR (47.4 %) and RATA (50 %) all had 'very easy' as the most selected option. Just like with the other conditions, RATA-JAR was never rated as 'very long' or 'very difficult'.

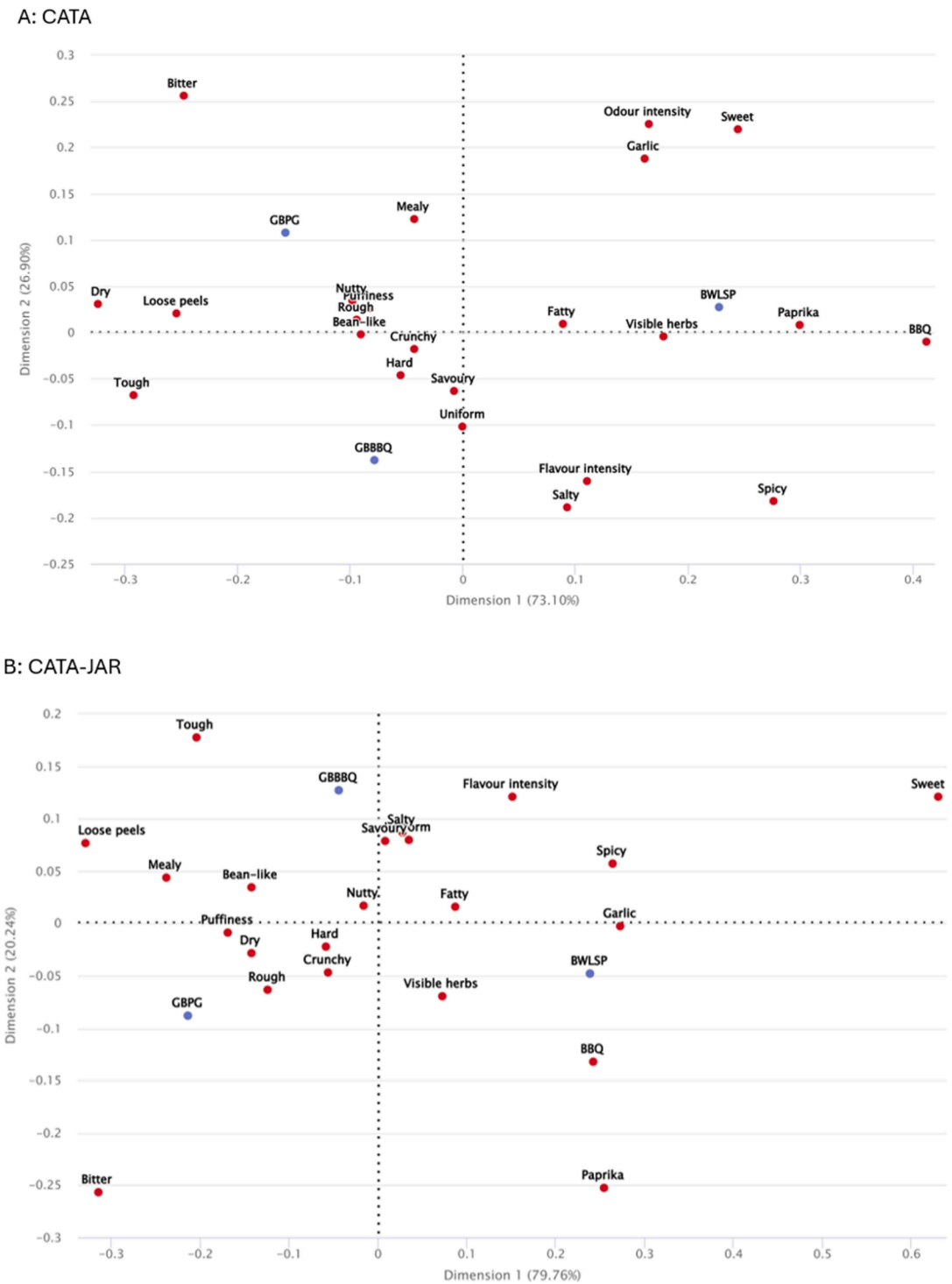

Figure 2. Correspondence analysis biplots for the four conditions: CATA (A), CATA-JAR (B), RATA (C) and RATA-JAR (D). In blue the visualization of the products GBPG, GBBBQ and BWLSP and in red the attributes.

4. Discussion

4.1. Attribute selection frequency

While selection frequency only increased between CATA and CATA-JAR (+5 % for BWLSP, +5.1 % for GBPG and +7.8 % for GBBBQ), and RATA and RATA-JAR slightly (+0.3 % for BWLSP, +0.5 % for GBBBQ and -3 % for GBPG), selection frequency showed a larger difference between CATA-JAR and RATA-JAR (25 %) and between CATA and RATA (28 %), which is similar to results in previous studies comparing CATA, CATA-JAR and/or RATA (Lee et al., 2021; Vidal et al., 2018). Vidal et al. (2018) explained this phenomenon by the change in the decision making process when the option 'low' is included. Response strategies differ between CATA and RATA, where in CATA participants only select the most applicable attributes instead of all the attributes that are present. Additionally, RATA might have a less natural evaluation by asking a analytical rating question (Bruzzone et al., 2011, 2015; Vidal et al., 2018). Higher selection frequency might make it harder for

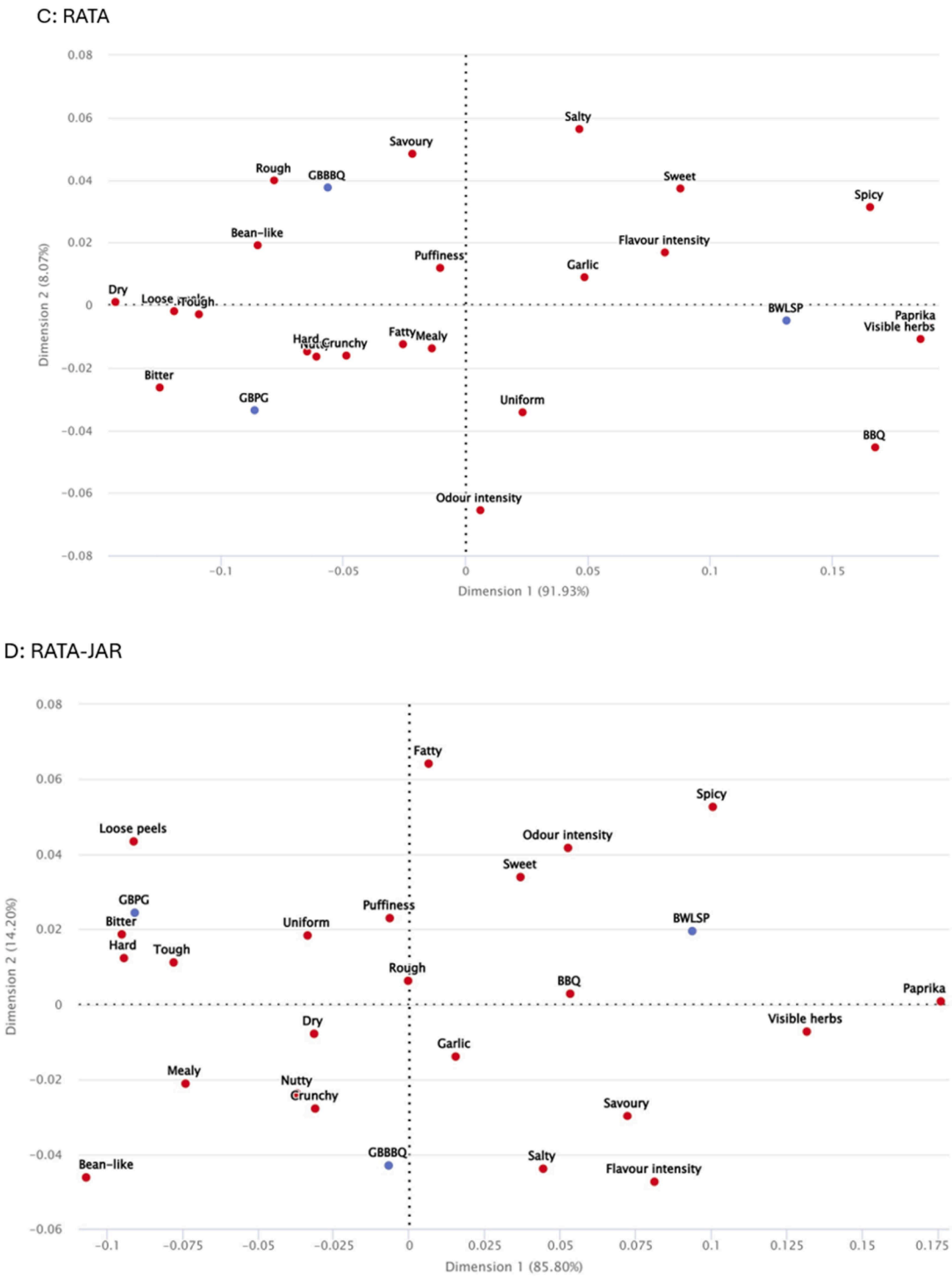

Figure 2. (*continued*).

product developers to know which attributes distinguish their products and therefore CATA-JAR might result in less selected attributes and a direction on how selected attributes are perceived from JAR, giving more precise input for product optimization or development compared to RATA(-JAR). The biplot of the CA of the four conditions showed similar dimensions and placement of products and attributes, so when only product profiles are needed without intensity or acceptability insights, CATA might be sufficient.

The research from Vidal et al. (2018) showed that RATA data should not be analysed as binary/CATA response as analysing RATA as binary data lead to a decrease in sample discriminations compared to CATA or RATA analysed as ordinal data. In our study, RATA analysed as binary data showed a lower percentage of attributes with a significant difference between products than when analysed as ordinal data (RATA 33.3 % versus 25 % and RATA-JAR 29.2 % versus 12 %). For RATA and RATA-JAR analysed as ordinal data slightly more attributes significantly differed between products compared to CATA (29.2 %), but less than CATA-JAR (37.5 %). Not all conditions showed significant differences in the same attributes. For example, bean-like showed only a significant difference between products in CATA-JAR, while differences in spiciness was only significant between products in RATA and RATA-JAR.

Table 3
Distribution of rating scores from the RATA and RATA-JAR conditions.

	GBPG		GBBBQ		BWLSP	
	RATA	RATA-JAR	RATA	RATA-JAR	RATA	RATA-JAR
Not present	23.6 %[a]	22.2 %[a]	15.5 %[a]	15.5 %[a]	18.1 %[a]	18.1 %[a]
Low	33.6 %[a]	37.5 %[a]	27.5 %[a]	37.7 %[a]	34.6 %[a]	35.0 %[a]
Medium	25.7 %[a]	20.6 %[a]	39.4 %[a]	30.1 %[a]	31.1 %[a]	29.9 %[a]
High	17.1 %[a]	19.7 %[a]	17.6 %[a]	16.7 %[a]	16.2 %[a]	17.1 %[a]

Different letters, [a-b] indicate significant differences between conditions within a product.

Table 4
Distribution of JAR data from the CATA-JAR and RATA-JAR.

	GBPG		GBBBQ		BWLSP	
	CATA-JAR	RATA-JAR	CATA-JAR	RATA-JAR	CATA-JAR	RATA-JAR
Too little	14.2 %	34.3 %[b]	11.4 %	23.6 %[b]	11.4 %	28.8 %[b]
Just About Right	63.8 %	41.0 %	77.3 %	66.0 %	77.3 %	56.8 %
Too much	22.1 %	24.8 %	11.4 %	10.4 %	11.4 %	14.4 %

Different letters, [a-b] indicate significant differences between conditions within a product.

4.2. Influence of adding JAR

Comparing the results on the JAR question between CATA-JAR and RATA-JAR, CATA-JAR appears to be closer to 'just about right' than RATA-JAR. While 'Just about right' and 'too much' did not show a significant difference between the two conditions, 'Too little' was selected significantly more often in RATA-JAR than in CATA-JAR for all three products. An explanation is that the 'low' attributes were not selected in CATA-JAR and therefore 'too little' was not chosen, whereas in RATA-JAR the selection frequency was higher and also 'low' intensity attributes were selected leading to more 'too little' selections in JAR. Perhaps, the consumer is adopting a more analytical mindset when including a rating question. The JAR scale suggests that there is a product that perfectly suits the consumers' preferences. This might give the idea that the product presented is not perfect, maybe even further from ideal than the respondent initially thought (Earthy et al., 1997; Popper et al., 2004). Because of this reasoning, participants may select the middle option ('medium' in the RATA question and 'just about right' in the JAR question) less frequently. It might also be that when participants indicated 'low' intensity in the rating-task, it influenced their choice of 'too little' in the JAR-task. However, this effect should be similar between 'high' and 'too much', but in our study selection of 'too much' was not significantly different between CATA-JAR and RATA-JAR. Because more attributes diverse from 'just about right' in RATA-JAR, it could be more difficult for product developers to know which attributes are critical to alter in further product development compared to CATA-JAR.

4.3. RATA and RATA-JAR

Distribution of RATA scores did not significantly differ between RATA and RATA-JAR. For the majority, but not all, attribute rating was correlated with JAR. A rating of 'low' for RATA does not equal a rating of 'too little' for JAR, just as 'medium' does not always mean 'just about right', and 'high' does not always align with 'too much'. If these ratings were directly corresponding to each other, the combination of RATA and JAR would offer little additional value. The potential benefit of RATA-JAR is indicating if the intensity of an attribute (low, medium, high) is experienced as positive or negative (JAR).

4.4. Length and difficulty

The level of difficulty of RATA-JAR significantly differed from the other three conditions, there were fewer selections for 'very easy' and more selections for 'a bit difficult'. Although the scores on length and difficulty were more negative compared to the other conditions, RATA-JAR was still mostly rated as easy/neutral with a good length. We were unable to compare the actual time it took to complete the test in each condition as all participants received the sample in each round at the same time. Participants were informed about this procedure, but were not informed about the different conditions.

4.5. Attribute selection

The selection of attributes that were easy to understand and suitable for all conditions was challenging. We determined- and confirmed understanding of attributes with a naïve consumer panel. However, not all attributes that work in CATA are suitable in RATA, for example 'strong flavour'; If 'strong flavour' is selected as present it does not make sense to rate it as 'low'. We only included sensory attributes, however one of the interesting and flexible options of CATA is that it is possible to include non-sensory terms. RATA-JAR seems less suitable when the objective is to also investigate differences in non-sensory attributes between products. For example, context of use might be a non-sensory category a developer wants to include, but item selection 'breakfast' does not make sense when asked about rating or JAR. An easy fix would be to apply separate categories, i.e. a sensory and non-sensory category and include RATA/JAR only for the sensory items. Consumers perception of the different tasks between sensory and non-sensory categories could be investigated.

4.6. Limitations and recommendations

The number of participants needed ranges from 25 to over 80 for CATA(-JAR) or RATA (Ares, Tárrega et al., 2014; Giacalone & Hedelund, 2016; Lee et al., 2021). We aimed to have 30 participants per condition, because of practical reasons and the exploratory nature of the study. Unfortunately, we had to exclude one of two test days, leading to a smaller sample size per condition and a lower power. With a larger sample size, differences between conditions would have probably been more pronounced. However, there were clear differences in selection frequency between conditions which is consistent with previous studies (Lee et al, 2021; Vidal et al, 2018). Additionally, the consumer panel used in the current study takes part in sensory tests once every two weeks on average and is familiar with sensory test procedures and terms, which improves performance as discussed by Ares & Varela (2017). Ares & Varela (2027) showed that several studies have shown that trained panels do not always perform better than consumer panels, and that if a trained panel performs better it may be due to familiarity with general test procedures. Another limitation is that this study only looked at the profiling of three bean snacks, therefore it remains unclear if the effect of the conditions on sensory profiles is similar for other types of products. We recommend exploring RATA-JAR in comparison to CATA and other variations further with a larger number of participants and products from different categories followed by an in-depth discussion on suitability with product developers from SMEs.

4.7. Conclusion

Sensory evaluation is an important part of product development. Especially for SMEs, while often dealing with limited resources, rapid sensory profiling techniques could lead to more useful sensory insights compared to suboptimal executed traditional methods. From these rapid methods, CATA, CATA-JAR, RATA and RATA-JAR are cost-effective

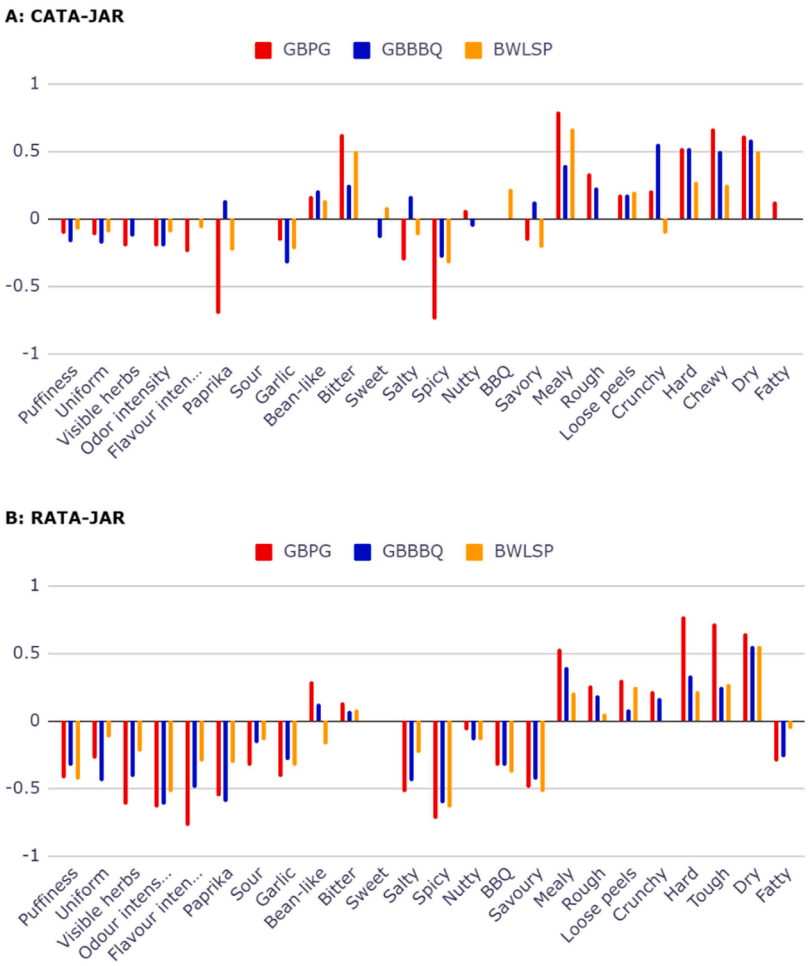

Figure 3. Visualisation of the JAR data for GBPG, GBBBQ and BWLSP with averages in bar charts for CATA-JAR (A) and RATA-JAR (B). Ranging from 'too little' (-1), just about right (0) to too much (+1).

ways to collect information with consumers on product characteristics. The additional benefit of RATA-JAR compared to CATA(-JAR), or RATA is questionable and depends on the test objective and products. As attribute selection frequency in RATA(-JAR) increased compared to CATA(-JAR), and compared to CATA-JAR, 'too low' was selected more frequently, this might distort the input for product development and make it less clear which are the important attributes to focus on. Compared to RATA, the addition of JAR might provide additional insights for product developers as JAR and rating were not always significantly correlated, meaning that for some attributes for example, a low rating does not necessarily mean 'too low' in JAR. In addition, length and difficulty was assessed similarly between the conditions. The most prominent reason to apply RATA-JAR would be when in product development input on intensity and the perception of that intensity (via JAR) is needed. We believe that often keeping it simple in practice, especially when resources are limited, by applying rating and/or JAR variations of CATA will be sufficient for several research objectives and make it easier for SMEs to obtain good and useful sensory insights in the product development process compared to the current applied methods.

Ethical statement

Participants gave informed consent after receiving general study information on procedures, voluntary nature of the study and data handling via the statement "I agree and would like to participate in this study". An affirmative reply was required to schedule a session to participate in the study. They were able to withdraw from the survey at any time without giving a reason. The products tested were safe for consumption.

CRediT authorship contribution statement

Vera van Stokkom: Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Safira Benguedda:** Writing – original draft, Visualization, Methodology, Investigation, Formal analysis. **Elise Pelgrim:** Writing – original draft, Methodology, Investigation. **Marieke Nijmeijer:** Writing – review & editing, Supervision, Methodology.

Declaration of competing interest

none

Acknowledgements

This study was funded by the Taskforce for Applied Research SIA, The Netherlands (project number HBOPD.2018.05.025). We thank GemBites for providing us with the bean snacks for the study. GemBites was not involved in the design, data collection, data analysis or reporting of this study. After analysis, main conclusions were shared with GemBites.

Appendix A. Average attribute selection frequency percentage for the four conditions per attritubte per product.*

	Product	CATA (n=18)	CATA-JAR (n=19)	RATA (n=18)	RATA-JAR (n=18)
Puffiness	GBPG	100 %$^{a;1}$	94.7 %$^{a,b;1}$	77.8 %$^{b;1}$	77.8 %$^{b;1}$
	GBBBQ	88.9 %$^{a;1}$	89.5 %$^{a;1,2}$	94.4 %$^{a;1}$	83.3 %$^{a;1}$
	BWLSP	83.3 %$^{a,b;1}$	68.4 %$^{b;2}$	94.4 %$^{a;1}$	88.9 %$^{a,b;1}$
Uniform	GBPG	44.4 %$^{a;1}$	42.1 %$^{a;1}$	94.4 %$^{b;1}$	100 %$^{b;1}$
	GBBBQ	55.6 %$^{a;1}$	57.9 %$^{a;1}$	88.9 %$^{b;1}$	100 %$^{b;1}$
	BWLSP	50.0 %$^{a;1}$	52.6 %$^{a;1}$	100 %$^{b;1}$	88.9 %$^{b;1}$
Visible herbs	GBPG	61.1 %$^{a;1}$	78.9 %$^{a;1}$	77.8 %$^{a;1}$	88.9 %$^{a;1}$
	GBBBQ	66.7 %$^{a;1}$	78.9 %$^{a,b;1}$	94.4 %$^{b;2}$	94.4 %$^{b;2}$
	BWLSP	94.4 %$^{a;2}$	100 %$^{a;2}$	100 %$^{a;2}$	100 %$^{a;3}$
High odour intensity	GBPG	50.0 %$^{a;1,2}$	52.6 %$^{a;1}$	88.9 %$^{b;1}$	77.8 %$^{a,b;1}$
	GBBBQ	27.8 %$^{a;1}$	52.6 %$^{a;1}$	83.3 %$^{b;1}$	88.9 %$^{b;1}$
	BWLSP	61.1 %$^{a;2}$	52.6 %$^{a;1}$	94.4 %$^{b;1}$	94.4 %$^{b;1}$
High flavour intensity	GBPG	38.9 %$^{a;1}$	42.1 %$^{a;1}$	94.4 %$^{b;1}$	100 %$^{b;1}$
	GBBBQ	61.1 %$^{a;1}$	73.7 %$^{a;2}$	100 %$^{b;\ 1}$	100 %$^{b;1}$
	BWLSP	61.1 %$^{a,b;1}$	73.7 %$^{a;2}$	100 %$^{b,c;\ 2}$	94.4 %$^{c;2}$
Paprika	GBPG	38.9 %$^{a;1}$	52.6 %$^{a;1}$	66.7 %$^{a;1}$	50.0 %$^{a;1}$
	GBBBQ	44.4 %$^{a;1}$	36.8 %$^{a,b;1}$	77.8 %$^{b;1}$	83.3 %$^{b;2}$
	BWLSP	77.8 %$^{a;2}$	89.5 %$^{a,b;2}$	100 %$^{b;2}$	88.9 %$^{a,b;3}$
Sour	GBPG	5.6 %$^{a;1}$	5.3 %$^{a;1}$	50.0 %$^{b;1}$	33.3 %$^{b;1}$
	GBBBQ	5.6 %$^{a;1}$	10.5 %$^{a,b;1}$	44.4 %$^{c;1}$	33.3 %$^{b,c;1}$
	BWLSP	0.0 %$^{a;1}$	10.5 %$^{a;1}$	44.4 %$^{b;1}$	38.9 %$^{b;1}$
Garlic	GBPG	44.4 %$^{a,\ b;1}$	31.6 %$^{a;1}$	77.8 %$^{c;1}$	66.7 %$^{b,\ c;1}$
	GBBBQ	27.8 %$^{a;1}$	47.4 %$^{a,b;1,2}$	77.8 %$^{b;1}$	77.8 %$^{b;1}$
	BWLSP	55.6 %$^{a;1}$	68.4 %$^{a;2}$	94.4 %$^{b;1}$	83.3 %$^{a,b;1}$
Bean-like	GBPG	100.0 %$^{a;1}$	94.7 %$^{a;1,2}$	94.4 %$^{a;1}$	94.4 %$^{a;1}$
	GBBBQ	94.4 %$^{a;1}$	100 %$^{a;1}$	94.4 %$^{a;1}$	88.9 %$^{a;1}$
	BWLSP	83.3 %$^{a;1}$	73.7 %$^{a;2}$	83.3 %$^{a;1}$	94.4 %$^{a;2}$
Bitter	GBPG	22.2 %$^{a;1}$	42.1 %$^{a,b;1}$	55.6 %$^{b,\ c;1}$	77.8 %$^{c;1}$
	GBBBQ	11.1 %$^{a;1}$	21.1 %$^{a;1}$	66.7 %$^{b;1}$	72.2 %$^{b;1}$
	BWLSP	11.1 %$^{a;1}$	21.1 %$^{a,b;1}$	50.0 %$^{b,c;1}$	66.7 %$^{c;1}$
Sweet	GBPG	38.9 %$^{a;1,2}$	5.3 %$^{b;1}$	50.0 %$^{a;1}$	55.6 %$^{a;1}$
	GBBBQ	22.2 %$^{a;1}$	36.8 %$^{a,b;2}$	61.1 %$^{b;1,2}$	50.0 %$^{a,b;1}$
	BWLSP	55.6 %$^{a;2}$	63.2 %$^{a;2}$	77.8 %$^{a;2}$	66.7 %$^{a;1}$
Salty	GBPG	50.0 %$^{a;1}$	68.4 %$^{a;1}$	72.2 %$^{a,\ b;1}$	94.4 %$^{b;1}$
	GBBBQ	83.3 %$^{a;2}$	94.7 %$^{a;2}$	94.4 %$^{a;1,2}$	100 %$^{a;1}$
	BWLSP	77.8 %$^{a;1,2}$	84.2 %$^{a;1,2}$	94.4 %$^{a;2}$	94.4 %$^{a;1}$
Spicy	GBPG	11.1 %$^{a;1}$	21.1 %$^{a;1}$	33.3 %$^{a,b;1}$	61.1 %$^{b;1}$
	GBBBQ	22.2 %$^{a;1}$	36.8 %$^{a;1}$	55.6 %$^{a,b;2}$	72.2 %$^{b;1}$
	BWLSP	27.8 %$^{a;1}$	47.4 %$^{a;1}$	83.3 %$^{b;2}$	94.4 %$^{b;2}$
Nutty	GBPG	83.3 %$^{a;1}$	78.9 %$^{a;1}$	94.4 %$^{a;1}$	83.3 %$^{a;1}$
	GBBBQ	72.2 %$^{a;1}$	89.5 %$^{a;1}$	100 %$^{a;1}$	66.7 %$^{a;1}$
	BWLSP	66.7 %$^{a;1}$	84.2 %$^{a;1}$	88.9 %$^{a;1}$	77.8 %$^{a;1}$
BBQ	GBPG	16.7 %$^{a;1}$	26.3 %$^{a,b;1}$	44.4 %$^{a,b;1}$	50.0 %$^{b;1}$
	GBBBQ	22.2 %$^{a;1}$	26.3 %$^{a;1}$	55.6 %$^{c;1}$	66.7 %$^{b;1}$
	BWLSP	44.4 %$^{a;1}$	47.4 %$^{a;1}$	94.4 %$^{b;2}$	72.2 %$^{a,b;1}$
Savoury	GBPG	83.3 %$^{a;1}$	63.2 %$^{a;1}$	88.9 %$^{a;1}$	77.8 %$^{a;1}$
	GBBBQ	94.4 %$^{a;1}$	84.2 %$^{a,b;1}$	94.4 %$^{b;1}$	88.9 %$^{a;2}$
	BWLSP	88.9 %$^{a;1}$	73.3 %$^{a;1}$	94.4 %$^{a;1}$	94.4 %$^{a;2}$
Mealy	GBPG	38.9 %$^{a;1}$	26.3 %$^{a;1}$	77.8 %$^{b;1}$	72.2 %$^{b;1}$
	GBBBQ	27.8 %$^{a;1}$	26.3 %$^{a;1}$	83.3 %$^{b;1}$	83.3 %$^{b;1}$
	BWLSP	33.3 %$^{a;1}$	15.8 %$^{a;1}$	83.3 %$^{b;1}$	77.8 %$^{b;1}$
Rough	GBPG	61.1 %$^{a;1}$	78.9 %$^{a;1}$	83.3 %$^{a;1}$	83.3 %$^{a;1}$
	GBBBQ	55.6 %$^{a;1}$	68.4 %$^{a,b;1}$	88.9 %$^{a,b;1}$	88.9 %$^{b;1}$
	BWLSP	50.0 %$^{a;1}$	63.2 %$^{a;1}$	72.2 %$^{a,b;1}$	94.4 %$^{b;1}$
Loose peels	GBPG	72.2 %$^{a;1}$	57.9 %$^{a;1}$	83.3 %$^{a;1}$	72.2 %$^{a;1}$
	GBBBQ	61.1 %$^{a,b;1,2}$	57.9 %$^{b;1}$	88.9 %$^{a;1}$	66.7 %$^{a,b;1}$
	BWLSP	38.9 %$^{a,b;2}$	26.3 %$^{b;2}$	77.8 %$^{c,a;1}$	66.7 %$^{c;1}$
Crunchy	GBPG	100 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;1}$
	GBBBQ	100 %$^{a;1}$	94.7 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;1}$
	BWLSP	94.4 %$^{a;1}$	94.7 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;1}$
Hard	GBPG	88.9 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;1}$
	GBBBQ	94.4 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;2}$
	BWLSP	83.3 %$^{a;1}$	94.7 %$^{a;1}$	100 %$^{a;1}$	100 %$^{a;2}$
Tough	GBPG	22.2 %$^{a;1}$	31.6 %$^{a,b;1}$	50.0 %$^{a,\ b;1}$	61.1 %$^{b;1}$
	GBBBQ	22.2 %$^{a;1}$	42.1 %$^{a,b;1}$	50.0 %$^{c;1}$	66.7 %$^{b;1}$
	BWLSP	11.1 %$^{a;1}$	21.1 %$^{a;1}$	38.9 %$^{a,b;1}$	61.1 %$^{b;1}$
Dry	GBPG	88.9 %$^{a,b;1}$	68.4 %$^{b;1}$	88.9 %$^{a,b;1}$	94.4 %$^{a;1}$
	GBBBQ	72.2 %$^{a;1}$	63.2 %$^{a;1}$	94.4 %$^{a;1}$	100 %$^{b;1}$
	BWLSP	38.9 %$^{a;2}$	52.6 %$^{a;1}$	83.3 %$^{b;2}$	100 %$^{b;1}$
Fatty	GBPG	22.2 %$^{a;1}$	42.1 %$^{a;1}$	77.8 %$^{b;1}$	94.4 %$^{b;1}$
	GBBBQ	22.2 %$^{a;1}$	52.6 %$^{a;1}$	77.8 %$^{b;1}$	83.3 %$^{b;1}$
	BWLSP	27.8 %$^{a;1}$	57.9 %$^{a,b;1}$	72.2 %$^{b,c;1}$	88.9 %$^{c;1}$

* Different letters$^{a\text{-}c}$ indicate significant differences ($p<0.05$) between conditions per attribute for each product based on dichotomous data for all conditions. Different numbers $^{1\text{-}3}$ indicate significant differences ($p<0.05$) between products within a condition based on dichotomous data for CATA and CATA-JAR and ordinal data for RATA and RATA-JAR.

Appendix B. Correlation coefficients between rating and JAR in RATA-JAR (N=18). (): number of participants that selected the attribute.*

Attribute	GBPG	GBBBQ	BWLSP
Puffiness	0.227 (14)	0.461 (15)	0.276 (16)
Uniform	0.030 (18)	0.775* (18)	0.452 (16)
Visible herbs	0.458 (16)	0.525* (17)	0.475* (18)
Odour intensity	0.897* (14)	0.733* (16)	0.692* ((17)
Flavour intensity	0.692* (18)	0.757* (18)	0.757* (17)
Paprika	0.804* (9)	0.600* (15)	0.600* (16)
Sour	0.316 (6)	0.316 (6)	0.643 (8)
Garlic	0.804* (12)	0.673* (14)	0.627* (15)
Bean-like	0.744* (17)	0.738* (16)	0.334 (17)
Bitter	0.826* (14)	0.651* (13)	0.174 (12)
Sweet	0.745* (10)	0.500 (9)	0.545 (12)
Salty	0.689* (17)	0.718* (18)	0.509* (17)
Spicy	0.542 (11)	0.539 (13)	0.861* (17)
Nutty	0.595* (15)	0.669* (14)	0.670* (14)
BBQ	0.857* (9)	0.754* (12)	0.887* (13)
Savoury	1.000* (14)	0.648* (16)	0.822* (17)
Mealy	0.530 (13)	0.656* (15)	0.688* (14)
Rough	0.538* (15)	0.453 (16)	0.593* (17)
Loose peels	0.641* (13)	0.716* (12)	0.991* (12)
Crunchy	0.198 (18)	0.200 (18)	0.210 (18)
Hard	0.478* (18)	0.707* (18)	0.676* (18)
Tough	0.856* (11)	0.730* (12)	0.856* (11)
Dry	0.698* (17)	0.517* (18)	0.657* (18)
Fatty	0.477 (17)	0.364 (15)	0.293 (16)

*Indicates significance ($p<0.05$).

Data availability

Data will be made available on request.

References

Ares, G., Bruzzone, F., Vidal, L., Cadena, R. S., Giménez, A., Pineau, B., Hunter, D. C., Paisley, A. G., & Jaeger, S. R. (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. https://doi.org/10.1016/j.foodqual.2014.03.006

Ares, G., Giménez, A., Barreiro, C., & Gámbaro, A. (2010). Use of an open-ended question to identify drivers of liking of milk desserts. Comparison with preference mapping techniques. *Food Quality and Preference, 21*(3), 286–294. https://doi.org/10.1016/j.foodqual.2009.05.006

Ares, G., & Jaeger, S. R. (2013). Check-all-that-apply questions: Influence of attribute order on sensory product characterization. *Food Quality and Preference, 28*(1), 141–153. https://doi.org/10.1016/j.foodqual.2012.08.016

Ares, G., & Jaeger, S. R. (2015). Examination of sensory product characterization bias when check-all-that-apply (CATA) questions are used concurrently with hedonic assessments. *Food Quality and Preference, 40*, 199–208. https://doi.org/10.1016/j.foodqual.2014.10.004

Ares, G., Tárrega, A., Izquierdo, L., & Jaeger, S. R. (2014). Investigation of the number of consumers necessary to obtain stable sample and descriptor configurations from check-all-that-apply (CATA) questions. *Food Quality and Preference, 31*, 135–141. https://doi.org/10.1016/j.foodqual.2013.08.012

Ares, G., & Varela, P. (2017). Trained vs. consumer panels for analytical testing: Fueling a long lasting debate in the field. *Food Quality and Preference, 61*, 79–86. https://doi.org/10.1016/j.foodqual.2016.10.006

Bruzzone, F., Ares, G., & Giménez, A. (2011). Consumers' texture perception of milk desserts. ii - comparison with trained assessors' data. *Journal of Texture Studies, 43*(3), 214–226. https://doi.org/10.1111/j.1745-4603.2011.00332.x

Bruzzone, F., Vidal, L., Antúnez, L., Giménez, A., Deliza, R., & Ares, G. (2015). Comparison of intensity scales and CATA questions in new product development: Sensory characterisation and directions for product reformulation of milk desserts. *Food Quality and Preference, 44*, 183–193. https://doi.org/10.1016/j.foodqual.2015.04.017

Cooper, R. G. (2010). *The Stage-Gate Idea to Launch System*. Wiley International Encyclopedia of Marketing. https://doi.org/10.1002/9781444316568.wiem05014

Delarue, J. (2015). 1 - The use of rapid sensory methods in R&D and research: an introduction. In J. Delarue, J. B. Lawlor, & M. Rogeaux (Eds.), *Rapid Sensory Profiling Techniques* (pp. 3–25). Woodhead Publishing. https://doi.org/10.1533/9781782422587.1.3.

Delarue, J., Lawlor, B., & Rogeaux, M. (2014). Rapid sensory profiling techniques and related methods: applications in new product development and consumer research. rapid sensory profiling techniques. *Elsevier Science & Technology*. https://doi.org/10.1016/B978-1-78242-248-8.50025-3

Dooley, L., Lee, Y., & Meullenet, J. (2010). The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. *Food Quality and Preference, 21*(4), 394–401. https://doi.org/10.1016/j.foodqual.2009.10.002

Earthy, P. J., MacFIE, H. J. H., & Hedderley, D. (1997). Effect of question order on sensory perception and preference in central location trials. *Journal of Sensory Studies, 12*(3), 215–237. https://doi.org/10.1111/j.1745-459X.1997.tb00064.x

GemBites. (2024). Gezond snacken begint bij GemBites! LinkedIn. Retrieved November https://www.linkedin.com/company/gembites/posts/?feedView=all. Retrieved November.

Giacalone, D., & Hedelund, P. I. (2016). Rate-all-that-apply (RATA) with semi-trained assessors: an investigation of the method reproducibility at assessor-, attribute- and panel-level. *Food Quality and Preference, 51*, 65–71. https://doi.org/10.1016/j.foodqual.2016.02.017

Howieson, J., Lawley, M., & Selen, W. (2014). New product development in small food enterprises. *Journal of New Business Ideas and Trends, 12*(1), 11. https://www.proquest.com/docview/1548331295.

Jaeger, S. R., & Ares, G. (2015). RATA questions are not likely to bias hedonic scores. *Food Quality and Preference, 44*, 157–161. https://doi.org/10.1016/j.foodqual.2015.04.011

Jaeger, S. R., Chheang, S. L., Jin, D., Ryan, G. S., & Ares, G. (2023). How do CATA questions work? Relationship between likelihood of selecting a term and perceived attribute intensity. *Journal of Sensory Studies, 38*(4). https://doi.org/10.1111/joss.12833. n/a.

Jaeger, S. R., Hunter, D. C., Kam, K., Beresford, M. K., Jin, D., Paisley, A. G., Chheang, S. L., Roigard, C. M., & Ares, G. (2015). The concurrent use of JAR and CATA questions in hedonic scaling is unlikely to cause hedonic bias, but may increase product discrimination. *Food Quality and Preference, 44*, 70–74. https://doi.org/10.1016/j.foodqual.2015.04.001

Kim, M., Heo, J., & Kwak, H. S. (2023). Comparison of sensory profiles by two different check-all-that-apply (CATA) terms developed from trained panelists and naïve consumers. *Food Quality and Preference, 109*, Article 104902. https://doi.org/10.1016/j.foodqual.2023.104902

Lee, S., Kwak, H. S., Kim, S. S., & Lee, Y. (2021). Combination of the Check-All-That-Apply (CATA) Method and Just-About-Right (JAR) Scale to Evaluate Korean Traditional Rice Wine (Yakju). *Foods (Basel, Switzerland), 10*(8), 1895. https://doi.org/10.3390/foods10081895

Pelgrim, E. (2022). *Development of CATA-JAR as Novel Sensory Method for Small and Medium Sized Enterprises*. Wageningen University & Research. Unpublished manuscript.

Pineau, N., Girardi, A., Lacoste Gregorutti, C., Fillion, L., & Labbe, D. (2022). Comparison of RATA, CATA, sorting and Napping® as rapid alternatives to sensory profiling in a food industry environment. *Food Research International, 158*, Article 111467. https://doi.org/10.1016/j.foodres.2022.111467

Popper, R., Rosenstock, W., Schraidt, M., & Kroll, B. J. (2004). The effect of attribute questions on overall liking ratings. *Food Quality and Preference, 15*(7), 853–858. https://doi.org/10.1016/j.foodqual.2003.12.004

Świąder, K., & Marczewska, M. (2021). Trends of using sensory evaluation in new product development in the food industry in countries that belong to the EIT regional innovation scheme. *Foods, 10*(2), 446. https://doi.org/10.3390/foods10020446

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. a review of novel methods for product characterization. *Food Research International, 48*(2), 893–908. https://doi.org/10.1016/j.foodres.2012.06.037

Vidal, L., Ares, G., Hedderley, D. I., Meyners, M., & Jaeger, S. R. (2018). Comparison of rate-all-that-apply (RATA) and check-all-that-apply (CATA) questions across seven consumer studies. *Food Quality and Preference, 67*, 49–58. https://doi.org/10.1016/j.foodqual.2016.12.013

Wang, S., Ng, K. H., Yee, K. H., Tang, Y., Meng, R., & He, W. (2023). Comparison of pivot profile, CATA, and Pivot-CATA for the sensory profiling of instant black coffee. *Food Quality and Preference, 108*, Article 104858. https://doi.org/10.1016/j.foodqual.2023.104858