Ranking task versus intensity scales, what is the best method to evaluate odor-induced taste enhancement?

Christopher Aveline, Thierry Thomas-Danguin, Charlotte Sinding

Received: 21 September 2022 | Revised: 24 May 2023 | Accepted: 29 May 2023

DOI: 10.1111/joss.12857

ORIGINAL ARTICLE

Journal of Sensory Studies — WILEY

Ranking task versus intensity scales, what is the best method to evaluate odor-induced taste enhancement?

Christopher Aveline | Thierry Thomas-Danguin | Charlotte Sinding

Centre des Sciences du Goût et de l'Alimentation, INRAE, Université Bourgogne Franche-Comté, CNRS, Institut Agro Dijon, Dijon, France

Correspondence
Charlotte Sinding, INRAE – CSGA Centre des Sciences du Goût et de l'Alimentation – INRAE 17 rue Sully, 21000, Dijon, France.
Email: charlotte.sinding@inrae.fr

Funding information
ISITE-BFC, Grant/Award Number: ANR-15-IDEX-0003; Institut National de Recherche pour l'Agriculture, l'alimentation et l'Environnement (INRAE), Grant/Award Number: half-PhD-funding; Agence Nationale de la Recherche (ANR): AROMA, Grant/Award Number: ANR-20-CE21-0004-01

Abstract

Odorants in food or beverages might enhance the taste. This phenomenon is called Odor-Induced Taste Enhancement (OITE). OITE has been shown using taste-intensity visual analog scales (VAS). VAS is often criticized because it is prone to halo-dumping effects. We compared VAS and ranking methods to evaluate OITE in people living with normal-weight or obesity. Sweet apple juice and salty green-pea soup were spiked with either vanillin or bacon odorants to produce OITE. In the VAS experiment, saltiness, sweetness, sourness, bitterness and the global aroma intensities were evaluated for each base with and without the odorants. The ranking task consisted in ranking from the lowest to the highest sweetness/saltiness intensities, three references with increasing tastant concentration and a target solution with the lowest tastant concentration and the odorant. VAS highlighted OITE neither in the apple juice nor in the green-pea soup, in no group. The ranking method revealed an OITE in the green-pea soup in both groups, and only the group with obesity experienced OITE in the apple juice.

Practical Applications

The ranking task appears as an optimal method to highlight OITE and is sufficiently sensitive to demonstrate subtle differences related to participants' weight status. The ranking task is easy to perform, does not require training and does not imply a high number of participants to be statistically powered. Odor-Induced Taste Enhancement is a strategy that might be employed to significantly reduce salt, and sugar, while maintaining an acceptable taste intensity, at home as well as in food formulation by the industry.

1 | INTRODUCTION

Food consumption involves the release of odorants and tastants in the mouth. These molecules reach the gustatory and olfactory receptors localized in the mouth and nose, respectively. The activation of these receptors produces sensory signals processed in dedicated and well-separated brain areas, which result in taste and odor perception. The integration of odor and taste results in the formation of a cognitive constructs known as flavor, which serves as a mental representation of the food. The memorization of a configural food object as a flavor, that is, odor and taste joint encoding, allows recalling the flavor from the olfactory modality only, which often leads to an increase in taste perception. Indeed, Odor-Induced Taste Enhancement (OITE) is the result of a transfer of taste property to an odor, which can further produce the enhancement of the taste by the odor. According to the associative learning theory of Stevenson and Prescott (Stevenson et al., 1995; Stevenson et al., 1998), repeated co-exposure to odorants and tastants leads to an association of the odor and taste as a flavor.

Several studies showed OITE with different odorants in sweet (Barba et al., 2018), salty (Lawrence et al., 2009) or bitter food bases (Labbe et al., 2006; Niimi et al., 2014). These studies mostly used visual analog scales (VAS) to assess taste intensity. Some experiments evaluated only the attribute of interest, namely a single taste intensity (Clark & Lawless, 1994; Djordjevic et al., 2004), while others considered in addition other attributes (other tastes, pleasantness, etc.) related to the tested samples (Clark & Lawless, 1994; Labbe et al., 2007; Lawrence et al., 2009; Niimi et al., 2014). The VAS method has already been employed to compare flavor perception between people living with normal–weight or obesity (Proserpio et al., 2017). In this study, the participants evaluated on VAS the sweetness intensity of custard creams in which an odorant eliciting a butter odor was added. The results showed an increased perception of the sweetness in participants with obesity when the odorant was added which was not the case for people with normal–weight. Therefore, VAS may seem an adequate method to show OITE and even subtle differences of OITE between people of different weight statuses. However, while everybody agrees to define vanillin as sweet, OITE is often difficult to show. A number of published studies highlighted no OITE effects while expecting to see one (Fondberg et al., 2021; Prescott et al., 2004; Stevenson et al., 1995; Sinding et al., 2022; Welge-Lüssen et al., 2005). All these studies used VAS and it is very likely that a number of unpublished studies also did not succeed in showing OITE.

We may therefore question the VAS method to show OITE effects. Indeed, VAS are prone to several biases among which are the range bias, the end-effect and the centering bias. The range bias occurs when participants use different implicit references as anchors of the scale or use more or less of the rating scale than expected (Lawless & Heymann, 2010a). The end-effect occurs when the participants avoid using the extremities of the scale just in case they get a lower or stronger stimulus later (Rousseau, 2004). The centering bias, as explained by Lawless and Heymann (2010a) "arises when subjects become aware of the general level of stimulus intensity they are likely to encounter in an experiment and tend to match the center or midpoint of the stimulus range with the midpoint of the response scale." Halo-dumping has been reported as another bias related to the use of VAS for quantitative sensory analyses of complex—food—products (Clark & Lawless, 1994). Clark and Lawless indicated that a dumping effect occurs "when a salient attribute is not included on a ballot, then the opinion about that aspect of the product may be displaced into another, sometimes inappropriate scale." As a consequence, and when applied to the study of OITE, the absence of a scale to evaluate the odor intensity might be responsible for an overestimation of the taste intensity. The dumping effect can be controlled by introducing all the appropriate attributes characterizing the product (i.e., tastes, odors, textures, colors) which are evaluated on different scales. However, introducing several attributes and scales can lead to another bias, namely the halo effect. This bias occurs when "the overall impression of a product affects the ratings of liking and other attributes and as a result, well-liked products may receive favorable ratings of all attributes" (Clark & Lawless, 1994). Therefore, using VAS in quantitative sensory analysis requires the use of the minimum number of scales needed to capture all the salient sensory characteristics of the products, as recommended for instance in the Quantitative Descriptive Analysis® procedure (Lawless & Heymann, 2010b). All these biases are therefore confusing to reliably interpret taste intensity results in flavored products and might reduce the accuracy of VAS to measure OITE.

When applied to the study of OITE, the most important caveat of the VAS method is that it forces perception toward analytical processing, which may alter the unity assumption likely important in multisensory integration (Chen & Spence, 2017) and thus disrupts the configural approach necessary to perceive the flavor (Prescott, 1999). A configuration (e.g., savory-roasted-chicken flavor) is a unique perception integrating the perceptions of the components (odor + taste). The flavor is typically and only a configuration in that differentiation between odor and taste disrupts the configuration, which has for consequence that the flavor is not perceived (Prescott et al., 2004). The components of the flavor configuration can be made independent if an elemental approach is taken. Accordingly, the rating of multiple attributes (e.g., aroma, taste), required on the one hand to control for the dumping effect in the VAS method, forces on the other hand the analytical processing of the flavor and cancels the perceptual integration of odor and taste that produces flavor perception (Prescott et al., 2004). This consideration underlines the limit of using VAS to study OITE since configural flavor perception is at the basis of the OITE effect.

While many textbooks on sensory evaluation recommend to use scales because they have several advantages, among which a monadic evaluation of the samples and a continuous rating that provides quantitative measures (Lawless & Heymann, 2010c; Strigler et al., 2009), the ranking task is an interesting alternative to the use of scales. Ranking is based on a comparison procedure, in which subjects compare pairwise samples and rank them according to one attribute. The ranking is an extension of a paired comparison procedure with more than two samples, and can be applied to hedonic or intensity evaluations. Many advantages of the ranking procedure have been described, for example the easiness of the task even for untrained participants or the needless of reference or anchoring samples (Lawless & Heymann, 2010c; Strigler et al., 2009). Moreover, this method produces data which are easy to interpret (Strigler et al., 2009) with no need of sophisticated data processing (Lawless & Heymann, 2010c). In the case of OITE studies, which depends on the use of a configural processing to perceive the flavor, it is likely that the ranking task appears to fulfill this requirement, since no analytical strategy is engage so that flavor can be processed configurally. Indeed, former studies have shown that the ranking task (RT) was efficient to show OITE in water and complex solutions (Aveline et al., 2022; Labbe et al., 2007; Nguyen, 2000). However, a formal comparison between VAS and RT to assess OITE was never performed.

The OITE can be difficult to assess in experimental conditions due to its intrinsic configural nature. This peculiar sensory property questions the adequacy of classical sensory methodologies that force elemental processing of complex perceptions. The objective of the present study was therefore to determine which sensory methods would allow to identify and assess the Odor-Induced Taste Enhancement and whether the methods are accurate to show subtle differences between populations of different weight status. Indeed, in a former study we found that OITE was more effective in OB compared to NW (Aveline et al., 2022). We tested the hypothesis that a ranking task would be more efficient to show OITE than rating intensities on Visual Analog Scales. To further check this hypothesis, we compared the efficiency of both psychophysical methods to assess OITE differences between people living with obesity (OB) or normal weight (NW).

2 | MATERIALS AND METHODS

2.1 | Participants

In experiment 1 (VAS method), we recruited 38 normal-weight (NW: 20 < body mass index, BMI < 25 kg/m^2) and 44 volunteers living with obesity (OB: BMI > 30 kg/m^2). In experiment 2 (RT method), 43 normal-weight (NW: 20 < BMI < 25 kg/m^2) and 38 volunteers living with obesity participated (OB: BMI > 30 kg/m^2) (Table S1). We performed a power analysis with Gpower (v3.1.9.7) using the data acquired during a pre-test on 16 participants. The alpha and beta parameters were set up respectively at 5% and 80%. The analysis suggested a minimum of 40 participants to reach significance level for both VAS and RT experiments. In both experiments, the participants were scaled and weighed at the end of the first session. They should not eat, drink or smoke at least 1:30 h before the sessions. The experimental procedure was explained to each participant before recruitment and again before each test session. Participants signed an informed consent form to participate in the study. They received 10€ compensation for each hour spent performing evaluations. The study was conducted following the Helsinki Declaration and was approved by the ethics committee OUEST III #19.04.26 (ID RCB #2019-A00120-57).

2.2 | Solutions

A series of pre-tests was performed to determine (i) the bases and aromas to be used in both experiments (pre-test 1), (ii) the salt and sucrose concentrations of references needed for the second experiment (pre-test 2). In pre-test 1, groups of 3–6 expert participants within the CSGA lab evaluated 8 sweet and salty bases and 42 odorants. Two complex bases were selected, along with odorants that had the potential to increase sweetness/saltiness, while avoiding pungency, long-lasting in mouth, and extreme unpleasantness. These three criteria were implemented to prevent any disruptions during the ranking sessions. Pre-test 2 was conducted with a panel of 16 expert participants from the CSGA lab to validate the sugar/salt concentrations of the references. The references were created following two requirements regarding intensity levels: distinct enough to allow an easy discrimination, yet close enough to enable precise ranking of the target solution. In addition, the highest reference concentration was chosen to be close to commercially available products similar to the selected bases. The lowest reference was set as the half of the highest concentration, or lower, but had to be still easily detectable by all participants in the pre-tests. The intermediate reference was the determine as the median between the lowest and the highest concentrations. Participants performed 12 rankings with the selected references and part of the solutions with added odorants. At least 2/3 of the participants correctly ranked the references, validating the final choices. Ultimately, two complex bases and two odorants were chosen for the formal test.

The base solutions were apple juice (Aj, Apple juice 100% pure juice Carrefour Extra, Carrefour, France), green pea soup (Gp, baby food 4–6 months bio green pea, Bledina, France). The Aj base was constituted of apple juice, initially containing 10% (w/w) of sugar, and diluted at 4% (w/w) with water (Evian, France). Food grade vanillin (CAS: 121-33-5, Sigma Aldrich, France) was selected to be tested as sweetness enhancer. In experiment 1, a vanillin solution was prepared first at 0.4% (w/w) in water and then diluted in Aj at 0.02% (w/w). In experiment 2, vanillin, was dissolved in Aj at 0.03% (w/w). The Gp base was prepared by extracting the supernatant of the unsalted green-pea puree (baby food) after centrifugation (Beckman Coulter, 15,000 RPM, 20°C, 30 min, acceleration max, deceleration min). Then for both experiments, 0.25% (w/w) of sodium chloride (Sigma-Aldrich) was dissolved in the Gp supernatant. A food grade smoky bacon odorant composition (Bacon smoked flexarome, 880,501 FB542, Firmenich, Switzerland) was chosen as a potential saltiness enhancer for Gp. The bacon odorant composition was dissolved in Gp at 0.005% (w/w).

In experiment 1, the participants evaluated 2 solutions for sweetness (the sweet apple juice and the vanillin-added sweet apple juice) and 2 for saltiness (the salty green-pea soup and the bacon-added salty green-pea soup). For experiment 2, three reference solutions with increasing concentrations (S1–S3) of either sucrose (Béghin Say, Tereos, France) or sodium chloride (Cooper, France) were prepared for each base. A fourth solution containing the odorant diluted in the S1 solution was prepared as specified in Table 1.

2.3 | Sensory procedure

The data collection was monitored with Fizz software (Biosystèmes, Couternon, France). For both experiments, the solutions were delivered using spray bottles, which avoid orthonasal odor perception before in-mouth tasting. At the beginning of the session, participants were trained to use the spray bottles. In experiment 1, the participants received 4 solutions monadically in a counterbalanced order with an interstimulus interval of 30 s. They were instructed to spray two pulses of the solution in mouth, to swallow, and rate sweetness,

TABLE 1 Concentration (w/w) of odorant and tastant compounds in the bases for the visual analogue scale (VAS, experiment 1) and the ranking tasks (RT, experiment 2); adapted from Aveline et al. (2022).

Experiment	Beverage base	Odorant	S1	S2	S3	S1 + odorant
Experiment 1: VAS	Apple Juice (Aj)	Vanillin	4%	NA	NA	4% + 0.02%
	Green-pea soup (Gp)	Bacon	0.25%	NA	NA	0.25% + 0.005%
Experiment 2: RT	Apple Juice (Aj)	Vanillin	4%	6%	8%	4% + 0.03%
	Green-pea soup (Gp)	Bacon	0.25%	0.50%	0.75%	0.25% + 0.005%

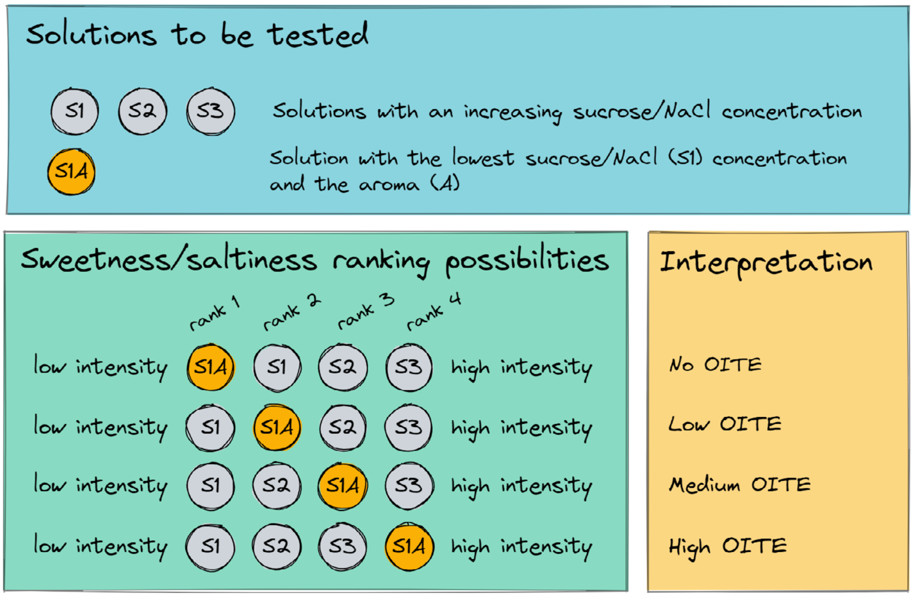

FIGURE 1 Overview of the ranking procedure to assess the odor-induced taste enhancement (OITE) from Aveline et al. (2022).

saltiness, sourness, bitterness and global aroma intensities on separate visual analog scales (VAS) presented on the same screen. The scales were labeled from left to right with "not intense" and "very intense." Participants could retest the solutions as much as necessary to evaluate all the attributes. When necessary, they could rinse their mouth with lukewarm Evian© water (40°C).

In experiment 2, the participants received four bottles (i.e., the three sweet or the three salty reference solutions (AjS1, AjS2, and AjS3 or GpS1, GpS2, and GpS3) and the target odorant-added solution (AjS1 + vanillin or GpS1 + bacon)). They were asked to rank the 4 bottles according to sweetness or saltiness intensity, from the lowest to the highest intensity (Figure 1). The ranking tasks' and solutions' orders were counterbalanced across participants. To perform the ranking task participants had to spray two pulses of each solution in the mouth, swallow and rank the bottles; ties were not allowed accordingly to guidelines provided by Cleaver (2018). When necessary, they could rinse their mouth with lukewarm Evian© water (40°C). After a preliminary ranking, they were instructed to confirm their choice by testing once again each solution from the lowest to the highest saltiness or sweetness intensity, before registering their final ranking. Reassessment of the samples is important to optimize the sensitivity of the ranking task (Cleaver, 2018). Between two ranking tasks, the participants had to rinse their mouth with lukewarm water and wait for 30 s.

At the end of experiment 1 and 2, the participants rated the pleasantness of the solutions on an unstructured linear scale, labeled from left to right with "not all pleasant" and "very pleasant." The samples were presented monadically in a counterbalanced order across participants. The participants had to rinse their mouth with lukewarm water between each sample.

2.4 | Data analysis

Data analysis was performed with R 3.5.1 (R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2020). For experiment 1, values of OITE were calculated as defined by Nasri et al. (2011). For each participant, the saltiness/sweetness intensity of the odorant-added solution was subtracted from the value of the same solution without the odorant. A bilateral t-test was performed to compare mean sweetness/saltiness OITE against 0. A significant t-test indicated an odor-induced sweetness/saltiness enhancement if OITE > 0, or an odor induced sweetness/saltiness suppression if OITE < 0. Moreover, we defined the strength of OITE according to the limit of the effect size obtained from the Cohen test: value < 0.2 SD (standard deviation): no OITE, 0.2 SD $<$ value < 0.5 SD: low OITE, 0.5 SD $<$ value < 0.8 SD: medium OITE, value > 0.8 SD: strong OITE. A t-test was carried out to compare the OITE between the NW and OB groups.

A Linear Mixed Effect model (LME) of ANOVA was applied to test the influence of the *solution* and *group* factors and their interaction on the raw *intensity* data of the target (i.e., for the sweet solutions: sweetness and for the salty solutions: saltiness) and nontarget sensory attributes (i.e., for sweet solutions: saltiness, bitterness, sourness and global aroma, and for salty solutions: sweetness, sourness, bitterness and global aroma); *participants* were modeled as a random factor. Pairwise Tukey post hoc tests with False Discovery Rate (FDR) correction for multiple comparisons were performed when a factor was significant ($p < .05$).

The data of the experiment 2 have been already published (Aveline et al., 2022) but were analyzed again here specifically in the context of the comparison of VAS and RT methods. As done in Aveline et al. (2022), a Wilcoxon test was carried out to check whether, references solutions were correctly ranked and therefore constituted a relevant scale of reference for taste intensity (i.e., S1 < S2 < S3). References were used as an "intensity scale" in the data analysis and as a quality control for data inclusion. An incorrect ranking of the references produces an invalid "reference scale" and the ranking of the OITE solution is inaccurate. Therefore, incorrect rankings of the references were removed to reduce noise in the data. When the ranking of these solutions was incorrect, the data of the participant were discarded. The percentage of discarded data was 32% on average in the apple juice (NW: 28%; OB: 37%) and 21% in the green-pea soup (NW: 14%; OB: 29%). Another Wilcoxon test was performed to compare the ranking of the solutions with odorant against the base solutions with different levels of sugar (or salt) (e.g., AjS1 vs. AjS1 + vanillin). When the solution with an odorant was perceived as sweeter/saltier than the same solution without the odorant, OITE was effective. A between-group Wilcoxon Mann Whitney test was performed to compare the ranking of the solution with odorant in each base. The pairwise Wilcoxon tests were corrected (FDR) for multiple comparisons; and were considered as significant when $p < .05$.

Specifically, here, a chi-squared test was used to compare the OITE distribution (number of participants who experienced OITE or no-OITE) between experiment 1 and 2 for all participants and by group (OB and NW). To facilitate the comprehension of the RT data (Aveline et al., 2022), the results were represented in terms of proportion of participants who ranked the samples in the first, second, third, or fourth position. The figures thus provide the distribution of rankings; however, the significance symbols (*) represent the difference of ranking between solutions (S1 vs. S1 + odorant) to show OITE effect. The ranking of the odorized sample in the second, third, and fourth positions revealed the OITE strength, corresponding respectively to a weak, medium, or strong taste enhancement relatively to the reference solutions S1, S2, and S3 (Figure 1).

A hunger index was calculated for each participant and each session by subtracting the index value at the end of the session and the value at the beginning of the session. We performed Student *t*-tests on mean index values to evaluate whether the hunger state changed during the sessions for each group, and to compare hunger scores at the begin of the experiment between the groups. A Linear Mixed Effect model (LME) of ANOVA was applied to test the effect of the factors *solution* and *group* on *pleasantness*; *participants* were modeled as a random factor. Pairwise Tukey post hoc tests with FDR corrections for multiple comparisons were performed when the factor *solution* was significant for each group ($p < .05$).

3 | RESULTS

3.1 | Experiment 1: Visual analog scales

In experiment 1, VAS were used to evaluate the intensity of sweetness, saltiness, sourness, bitterness and global aroma (Figures 2a, S1 and S2). In apple juice, the mean value of OITE was 0.41 in NW and −0.24 in OB. The *t*-test revealed no significant OITE based on Nasri's calculation (2011), neither in NW ($t[37] = 1.30$; $p = .20$) nor in OB ($t[43] = -1.10$; $p = .28$). Moreover, no difference was observed between the groups ($t[68] = 1.55$; $p = .13$). Based on the raw measures of intensities, the ANOVA showed no significant difference of sweetness between the solutions ($F[1,80] = 0.10$, $p = .76$) and between the groups ($F[1,80] = 0.02$, $p = .88$) (Figure S1). For the other taste attributes, no difference was found for saltiness ($F[1,80] = 0.15$, $p = .70$), sourness ($F[1,80] = 0.88$, $p = .35$), bitterness ($F[1,80] = 1.01$, $p = .32$) and global aroma intensity ($F[1,80] = 1.24$, $p = .27$). Moreover, no significant difference between groups was found for saltiness ($F[1,80] = 3.78$, $p = .06$), bitterness ($F[1,80] = 0.01$, $p = .91$) and global aroma intensity ($F[1,80] = 0.35$, $p = .56$). However, the NW group perceived the solutions as more sour than the OB group ($F[1,80] = 5.12$, $p = .03$).

In the green-pea soup, the mean value for OITE was 0.04 in NW and 0.15 in OB. The *t*-test revealed no significant OITE based on Nasri's calculation (2011), neither in NW ($t[37] = 0.11$; $p = .91$) nor in OB ($t[43] = 0.50$; $p = .62$). Moreover, no difference of OITE was observed between the groups ($t[78] = -0.25$; $p = .80$). Based on the raw measures of intensities, no difference of saltiness was found between the solutions (ANOVA, $F[1,80] = 0.19$, $p = .66$), but the NW group perceived the solutions as saltier than OB ($F[1,80] = 7.32$, $p = .008$) (Figure S1). For the other taste attributes, no significant difference was observed for sweetness ($F[1,80] = 0.77$, $p = .38$), sourness ($F[1,80] = 0.07$, $p = .79$), bitterness ($F[1,80] = 0.07$, $p = .80$) and global aroma intensity ($F[1,80] = 0.27$, $p = .60$). Moreover, no difference between groups was found for sweetness ($F[1,80] = 0.22$, $p = .64$), sourness ($F[1,80] = 0.19$, $p = .66$) the bitterness ($F[1,80] = 3.90$, $p = .05$) and global aroma intensity ($F(1,80) = 1.90$, $p = .17$).

3.2 | Experiment 2: Ranking task

In experiment 2, the participants were asked to rank four solutions according to sweetness or saltiness intensity. These results were published elsewhere (Aveline et al., 2022).

Concerning the apple juice, 75% of OB ranked AjS1 + vanillin at least in the second position, whose 25% in the third and 12% in the fourth position. By contrast, only 61% of NW ranked AjS1 + vanillin

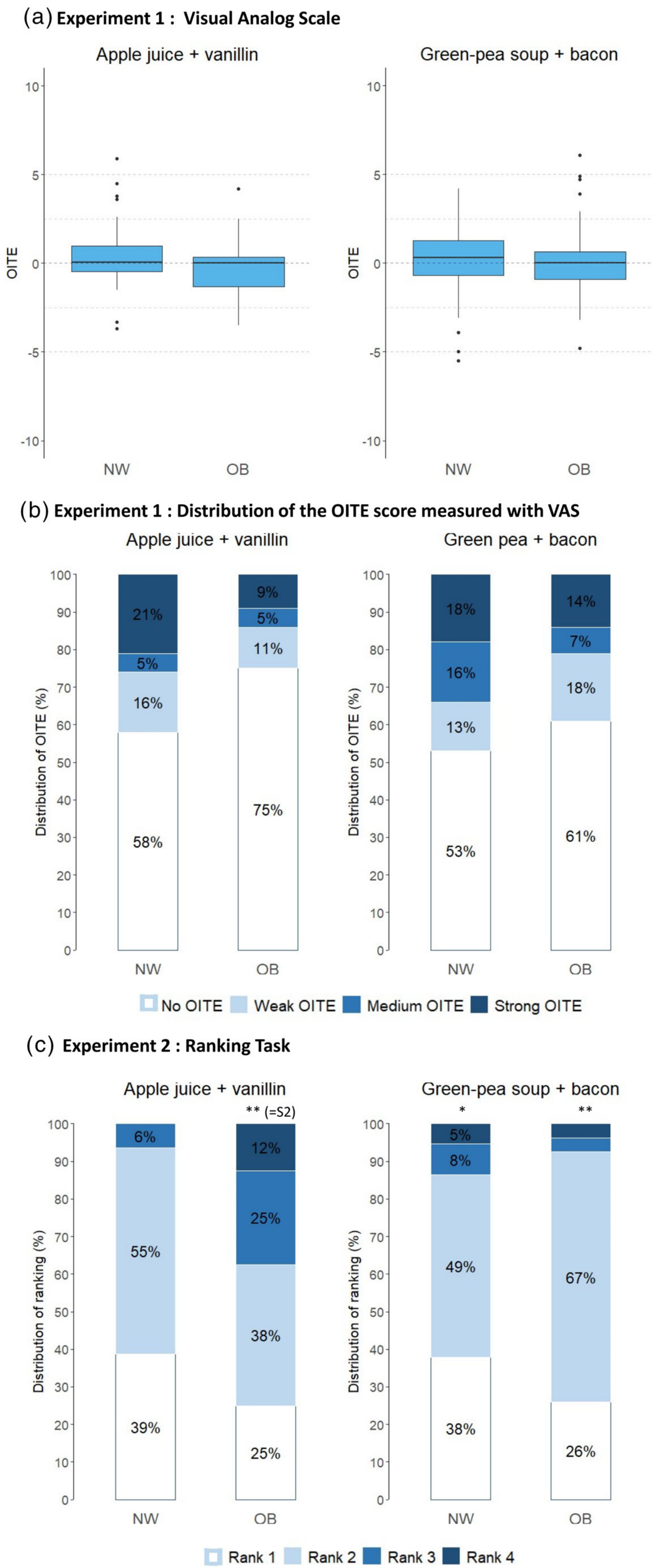

FIGURE 2 Legend on next page.

at least in the second position whose only 6% in the third position (Figure 2c). Therefore, AjS1 + vanillin was ranked as significantly higher than AjS1 by OB ($W = 252$, $p = .003$) but not by NW ($W = 316$, $p = .14$), which means that OB participants experienced OITE in apple juice but NW ones. Interestingly, OB perceived AjS1 + vanillin as sweet as AjS2 ($W = 108$, $p = .22$), corresponding to a medium level of OITE.

In the green-pea soup, 74% of OB ranked GpS1 + bacon at least in the second position whose 3.5% in third, and 3.5% in fourth position. By contrast, 62% of NW ranked GpS1 + bacon at least in the second position, whose 8% in third, and 5% in fourth (Figure 2c). Therefore, GpS1 + bacon was ranked as significantly saltier than GpS1 by OB ($W = 287$, $p = .009$) and by NW ($W = 472$, $p = .05$). However, NW and OB did not rank differently GpS1 + bacon ($W = 286$, $p = .62$), which means that participants of both groups experienced OITE in green-pea soup. Both groups evaluated the GpS1 + bacon as saltier than GpS1 but less salty than GpS2 (OB: $W = 33.5$, $p < .001$; NW: $W = 92$, $p < .001$), corresponding to a weak level of OITE.

3.3 | Comparison of experiment 1 and 2

The distribution of the OITE levels with VAS was calculated (Figure 2b). We compared the number of participants who presented OITE, or no OITE between the two methods (Figure 2b,c). In the apple juice, the results revealed a significant difference in the OITE distribution between the two methods, independently of the group; the RT showed that significantly more people experienced OITE than the VAS (χ^2 [1] = 14.25 $p < .001$). Within the OB group, the ranking method highlighted more people experiencing OITE (χ^2 [1] = 13.81, $p < .001$) than the VAS method, and no difference was found in the NW groups between methods (χ^2 [1] = 1.81, $p = .18$).

In the green-pea soup, the results revealed a difference in the OITE distribution between the two methods, the RT method also showed a higher number of participants who experienced OITE than with VAS (χ^2 [1] = 7.72, $p = .005$). The group analysis revealed that, in the OB group, the RT method showed more participants experiencing OITE than with VAS (χ^2 [1] = 7.06, $p = .008$). In the NW group no difference of distribution was found between the two methods (χ^2 [1] = 1.11, $p = .29$).

4 | DISCUSSION

The present study aimed to compare the Visual Analog Scale (VAS) and the Ranking Task (RT) methods to assess the Odor-Induced Taste Enhancement (OITE). Moreover, to check whether these methods could highlight subtle differences between participants with different weight statuses, OITE was measured in people living with obesity or normal weight. We hypothesized that RT would be more efficient than VAS to measure OITE and to show a difference of OITE between the two populations. In line with our hypothesis, our results showed that the RT was more efficient to demonstrate OITE. Moreover, only the RT method showed that people living with obesity experienced OITE in the apple juice in which vanillin was added, while it was not the case in the normal-weight group. In the green-pea soup, again only the RT method showed that both groups significantly experienced OITE with a bacon odor.

4.1 | The VAS method disrupts the OITE

Vanillin is the most used aroma in the food and flavor industry as an enhancer of sweetness perception (Banerjee & Chattopadhyay, 2019), and it has been repeatedly shown as an effective sweetness enhancer in sensory studies (cf. Spence, 2022, for a review). However, the results of our first experiment showed that the VAS method did not highlight OITE with vanillin. The VAS method has been traditionally used to show odor-induced taste enhancements (Frank & Byram, 1988; Lawrence et al., 2009; Nasri et al., 2011, 2013; Proserpio et al., 2017), but it is often questioned for its efficiency to measure OITE because it is prone to the dumping effect (Clark & Lawless, 1994) even if this effect was debated in the case of OITE (Prescott, 1999). Clark and Lawless recommended to include all the salient attributes in the measurement procedure to minimize the dumping effect. However, providing the different sensory attributes of a food or beverage engage panelists to process the elements of the flavor (e.g., the odor and taste separately) which has been shown to disrupt flavor perception (Prescott & Murphy, 2009). Conversely, engaging participants into a configural perception, by asking the participants to focus on the overall flavor resulted in OITE (Prescott et al., 2004). A configuration (e.g., a flavor such as a savory roasted chicken) is a unique perception integrating the perceptions of the

FIGURE 2 (a) Experiment 1: Sweet and salty OITE obtained from intensity ratings on visual analogue scales (VAS) in apple juice and green-pea soup model beverages. OITE was calculated as the difference between the sweet/salty solution with the odorant vs. the sweet/salty solution without the odorant. The vanillin odorant was used for apple juice, and a bacon odorant composition was used for the green-pea soup. The dotted lines represent different arbitrary levels of OITE: OITE < 0: No OITE, 0 < OITE < 2.5: weak OITE, 2.5 < OITE < 5: medium OITE or 5 < OITE: strong OITE. (b) Experiment 1: Distribution of the levels of OITE obtained from the VAS ratings. (c) Experiment 2: Distribution of the ranking scores for the apple juice base with the vanillin odorant, and for the green-pea soup with a bacon odorant composition. The color gradient corresponds to the distribution of the participants who ranked the target solution in the 1, 2, 3, or 4th position. Significant comparisons between the target solution (solution with odorant, S1 + odorant) and the reference (solution without odorant, S1) were indicated with ***: $p < .001$; **: $p < .01$; *: $p < .05$ (Wilcoxon test). The term in brackets indicates that the solution was perceived as sweet/salty as the S2 solution (adapted from Aveline et al. (2022)).

components (e.g., odor + taste). Therefore, the use of multiple VAS to evaluate the flavor may, on the one hand limit the dumping bias, but on the other hand, introduce another critical bias, namely the commitment to an elemental perceptual strategy that leads to disrupt the configural flavor and in turn the OITE.

4.2 | Validation of a comparative method to assess OITE

The ranking method is a comparative method in which participants perform a series of paired comparisons between the samples of a set. This method has already been employed successfully to highlight OITE (Aveline et al., 2022; Labbe et al., 2007; Nguyen, 2000). In the experimental design developed by Nguyen (2000), 10 solutions were used, whose 7 references with increasing concentration levels of sucrose in water and 3 of these solutions in which vanillin was added. In a first block, the participants ranked the odorless solutions from the lowest to the highest sweet taste intensity. In a second block, they ranked the 10 samples including the 7 references and the 3 samples with vanillin added. The authors compared the distribution of participants who shifted the solution with the vanillin to a higher rank compared with the matched reference without vanillin. They found that OITE occurred only when vanillin was mixed with lower sugar concentrations. This study was further implemented in two countries (France and Vietnam) and the authors tested lemon and vanilla odors as potential sweetness enhancers. The ranking task allowed to show differences between countries: vanilla increased more sweetness in France than Vietnam, while Lemon increased sweetness only in Vietnam. This result shows that RT is efficient to highlight subtle differences in OITE. Labbe et al. (2007) also used a RT task but with a lower number of references. They used four solutions containing the same concentration of sugar and in three of them they added an increasing concentration of one odorant. They tested 6 different odorants (Ethyl butyrate (fruity), benzaldehyde (almond), furaneol (caramel), isoamyl acetate (banana), maltol (caramel) and vanillin (vanilla)). The participants had to rank 4 samples according to their odor or sweetness intensity. They showed that several odorants significantly enhanced sweetness (Ethyl butyrate (fruity), benzaldehyde (almond), furaneol (caramel), isoamyl acetate (banana)) while maltol (caramel) and vanillin did not. This study confirms that OITE can be highlighted with a ranking method. In our study, we used a similar design as Nguyen and Labbe studies and further extended these results on the methodological aspects by comparing results between VAS and RT. While the VAS did not show OITE in any of the solutions tested, the RT method showed OITE in most but not all the solutions tested, which reflects the sensitivity of the method to discriminate taste enhancement induced by odors. Moreover, the simple ranking task developed here allowed to discriminate levels of OITE, corresponding to low, medium and high effects according to the comparison with the three reference solutions. It should be noticed that a ranking task should not exceed six samples to keep its effectiveness (Cleaver, 2018), as more samples would be too complex to rank and would results in more false ranking of the reference solutions.

Another study relying on a comparative task allowed to show OITE following an ABCDX experimental design (Wang et al., 2019). This method is based on a matching task between one target odorant-added sample (X) and 4 references made of no-odorant-added samples with increasing sucrose concentrations (ABCD). The subjects had to evaluate which reference sample (A, B, C, or D) was the most similar to the target sample (X). The results allowed the identification of OITE by comparing the distribution of people who matched the odor-added sample (X), which had the same sucrose level than reference A, to the higher sucrose content references. The data showed that both a sweet-congruent (vanilla) and a sweet-non-congruent (beef) odor can enhance sweetness in milk samples, but enhancement was greater for the congruent odor. Moreover, and in line with our findings, the authors found that OITE observed with the matching method was higher than OITE obtained via a scaling task performed in previous study from the same group (Wang et al., 2018). As discussed by the authors, it is likely that the matching method encouraged a configural cognitive strategy during evaluation. Nevertheless, a limit of the matching approach developed in this study is that the basis on which participants made their matches was unknown, especially because in complex close-to-real food samples several sensory dimensions might be perceived and used for the matching procedure (e.g., overall flavor or even pleasantness). The RT method we developed in our present study ruled out this issue since participants had to rank the samples according to sweetness or saltiness intensity. Such a methodological approach appears therefore optimal for the study of OITE as it focuses the evaluation on the target flavor but did not engage an analytical processing that would lead to take apart taste and smell and thus prevent configural flavor perception.

Furthermore, the RT allowed to discriminate OITE between OB and NW populations. While we observed odor-induced saltiness enhancement independently of weight status, odor-induced sweetness enhancement by vanillin in apple juice was observed only in OB. Indeed, our results showed that, among this population, 37% of the participants perceived the vanillin-added apple juice as sweeter than apple juice containing 33% more sugar. Difference of OITE between NW and OB has been reported in two studies from the same group using a VAS method (Proserpio et al., 2016, 2017). In these studies, only people living with obesity experienced odor-induced sweetness enhancement with a butter odorant added in a custard dessert. The butter odorant may be a cue for high-energy dense food and was found to also increased the vanilla odor intensity (Proserpio et al., 2016, 2017), which suggest that the odor-induced sweetness enhancement is likely due to a stronger odor-taste association in people living with obesity (Aveline et al., 2022). On the basis of these findings, the ranking task appears as a sound method to highlight subtle differences between populations.

4.3 | Benefits and limits of the ranking task to assess OITE

It has been shown that human perception is more comparative in nature (Lawless & Heymann 2010a, 2010b, 2010c). In other words, we perform better when evaluating products in relation to one another rather than in absolute terms. Cleaver (2018) made this point: "Even those of us without perfect pitch, for example, could identify which of two consecutive notes played on a piano was pitched higher, but would struggle when the interval is increased and other notes are interspersed." The motivation for ranking methodology is to exploit this natural mode of comparative perception and to present the products together so that they can be assessed against each other.

Conversely, ratings on VAS requires to rely on the assessor's memory to give repeatable absolute scores of intensities. Furthermore, VAS carries several implicit assumptions that cannot be met in group comparison (Bartoshuk et al., 2003). First, to compare groups of participants it is assume that the scale points are equally spaced in perceptual terms. Secondly, it is assumed that the intensity labels of the scale (anchor points) evoke the same experience on average in both groups. Ranking makes no such assumptions. Stevens (1958) made this point more than 40 years ago: "Mice may be called large or small, and so may elephants, and it is quite understandable when someone says it was a large mouse that ran up the trunk of the small elephant." Furthermore, ranking is likely more sensitive to detect subtle differences of intensities due to a more reliable relative positioning of the different samples as compared to VAS (Cleaver, 2018). Another benefit of RT is that the process of training of assessors is relatively short compared to what is usually required for the use of VAS, especially in the context of quantitative descriptive analysis. In our study, clear instructions on how to perform the task and use the software were provided before starting but no training was necessary. Essentially, assessors performing a ranking methodology need to understand the task and have a common agreed interpretation of the sensory attributes involved.

The ranking tasks also present some limits. First, the number of samples in one ranking set should not exceed 6 to 8, to limit fatigue and maintain the accuracy of the task (Cleaver, 2018). When odorants are used, long lasting odorants should be avoided to not produce carry over effects, and thus alter the evaluation of other samples in the set. Another limit is that RT is a semi quantitative technique that does not allow a continuous rating of intensities. To be able to semi-quantify the degree of enhancement, several reference solutions should be included in the sample set in addition to the target odor-added sample. Indeed, in our design, the references were used in the analysis as a "reference scale of intensities." Moreover, the success of the ranking task depends on the design of the taste levels of the reference solutions (i.e., the concentration in salt or sugar). The pre-requisite for the references was that they should be sufficiently distant to facilitate the correct ranking of the three levels of sugar/salt and close enough to allow a precise ranking of the sample including the target added odorant among the references. In the data analysis of the RT method, we excluded on average 16% of participants who did not rank correctly the three base solutions of each ranking (S1, S2, S3), which, on the one hand should strengthen the results, but on the other hand might increase the number of participants that should be recruited to participate in the study. Furthermore, this ranking procedure cannot be used in all types of experimental set-up, such as brain exploration experiments where participants cannot move during the brain activity recording and have a restricted period to evaluate the sample.

Finally, a limit of the present study is that the concentration of vanillin in the two experiments was not exactly the same (0.02% vs. 0.03%). Therefore, we could not exclude that the vanillin-induced sweetness enhancement observed in the second experiment was due to the slightly higher concentration of the odorant. Nevertheless, this is very unlikely because vanillin is an odorant with a low Stevens' slope (0.31, Devos et al., 2002), that reflects the evolution of the perceived intensity as a function of the odorant concentration as modeled by a power law. Thus, a low variation in vanillin concentration should not induce a significant variation in vanilla odor intensity. Indeed, in pre-testing the rankings of the two vanillin concentrations were not ranked differently (median van 200 ppm $= 2$, median van 300 ppm $= 2$; $W = 13$, $p = .22$). Finally, the results confirm that the ranking task is an effective and sufficiently powerful sensory method to highlight OITE.

5 | CONCLUSION

The ranking task with references of increasing concentrations of salt/sugar appears as an optimal sensory method to assess OITE. Furthermore, the RT method was sufficiently accurate to show differences of OITE between populations of different weight status. On the contrary, the VAS method did not show OITE for any of the solutions tested and did not show differences in perception between the two populations tested. The RT method is a comparative task easy to perform and more importantly that preserves the configural processing necessary to perceive the flavor while focusing the participants' evaluation on the target flavor. Conversely the VAS method, which requires the use of several attributes to avoid dumping effects, engages participants in analytical processing that disrupts OITE. Consequently, we recommend using the ranking task for future OITE sensory studies.

AUTHOR CONTRIBUTIONS

Christopher Aveline: Conceptualization, investigation, methodology, data curation, formal analysis, writing original draft. **Thierry Thomas-Danguin:** Conceptualization, investigation, methodology, formal analysis, validation, supervision, writing-review & editing. **Charlotte Sinding:** Conceptualization; funding acquisition; investigation; methodology; formal analysis; project administration; resources; software; supervision; validation; writing—review & editing.

ACKNOWLEDGMENTS

We would like to thank for their technical and logistic support on the experiment: Noëlle Béno, and Françoise Durey from the ChemoSens platform of CSGA, and Chantal Septier, Claire Follot, Gaia Maillard, Marianela Santoyo Zedillo, Marie Simon, Yue Ma and Anaïs Guilbert from the FFOPP team of CSGA. We would like to thank Firmenich (Geneva, Switzerland) for providing the bacon odorant composition.

FUNDING INFORMATION

This work was supported by the French "Investissements d'Avenir" program, project ISITE-BFC (contract ANR-15-IDEX-0003) awarded to Charlotte Sinding for the EATERS project, and by the INRAE department Transform for the PhD fellowship to Christopher Aveline. This work was also supported by a grant from the French National Research Agency awarded to Charlotte Sinding [Agence Nationale de la Recherche (ANR): AROMA project ANR-20-CE21-0004-01].

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Charlotte Sinding https://orcid.org/0000-0002-0887-4429

REFERENCES

Aveline, C., Leroy, C., Thomas-Danguin, T., & Sinding, C. (2022). Influence of obesity on saltiness and sweetness intensity enhancement by odors. *Food Quality and Preference*, *102*, 104685.

Banerjee, G., & Chattopadhyay, P. (2019). Vanillin biotechnology: The perspectives and future. *Journal of the Science of Food and Agriculture*, *99*, 499–506.

Barba, C., Beno, N., Guichard, E., & Thomas-Danguin, T. (2018). Selecting odorant compounds to enhance sweet flavor perception by gas chromatography/olfactometry-associated taste (GC/O-AT). *Food Chemistry*, *257*, 172–181.

Bartoshuk, L. M., Duffy, V. B., Fast, K., Green, B. G., Prutkin, J., & Snyder, D. J. (2003). Labeled scales (e.g., category, Likert, VAS) and invalid across-group comparisons: What we have learned from genetic variation in taste. *Food Quality and Preference*, *14*, 125–138.

Chen, Y. C., & Spence, C. (2017). Assessing the role of the "unity assumption" on multisensory integration: A review. *Frontiers in Psychology*, *8*, 1–22.

Clark, C. C., & Lawless, H. T. (1994). Limiting response alternatives in time-intensity scaling: An examination of the halo-dumping effect. *Chemical Senses*, *19*, 583–594.

Cleaver, G. (2018). Ranking and rank-rating. In S. E. Kemp, J. Hort, & T. Hollowood (Eds.), *Descriptive analysis in sensory evaluation* (pp. 447–491). John Wiley & Sons, Ltd.

Devos, M., Rouault, J., & Laffort, P. (2002). Standardized olfactory power law exponents. Dijon, France.

Djordjevic, J., Zatorre, R. J., & Jones-Gotman, M. (2004). Odor-induced changes in taste perception. *Experimental Brain Research*, *159*, 405–408.

Fondberg, R., Lundström, J. N., & Seubert, J. (2021). Odor-taste interactions in food perception: Exposure protocol shows no effects of associative learning. *Chemical Senses*, *46*, 1–14.

Frank, R. A., & Byram, J. (1988). Taste-smell interactions are tastant and odorant dependent. *Chemical Senses*, *13*, 445–455.

Labbe, D., Damevin, L., Vaccher, C., Morgenegg, C., & Martin, N. (2006). Modulation of perceived taste by olfaction in familiar and unfamiliar beverages. *Food Quality and Preference*, *17*, 582–589.

Labbe, D., Rytz, A., Morgenegg, C., Ali, S., & Martin, N. (2007). Subthreshold olfactory stimulation can enhance sweetness. *Chemical Senses*, *32*, 205–214.

Lawless, H. T., & Heymann, H. (2010a). Context effects and biases in sensory judgment. In H. T. Lawless & H. Heymann (Eds.), *Sensory evaluation of food* (pp. 203–225). Springer.

Lawless, H. T., & Heymann, H. (2010b). Descriptive analysis. In H. T. Lawless & H. Heymann (Eds.), *Sensory evaluation of food* (pp. 227–257). Springer.

Lawless, H. T., & Heymann, H. (2010c). Scaling. In H. T. Lawless & H. Heymann (Eds.), *Sensory evaluation of food* (pp. 149–177). Springer.

Lawrence, G., Salles, C., Septier, C., Busch, J., & Thomas-Danguin, T. (2009). Odour-taste interactions: A way to enhance saltiness in low-salt content solutions. *Food Quality and Preference*, *20*, 241–248.

Nasri, N., Beno, N., Septier, C., Salles, C., & Thomas-Danguin, T. (2011). Cross-modal interactions between taste and smell: Odour-induced saltiness enhancement depends on salt level. *Food Quality and Preference*, *22*, 678–682.

Nasri, N., Septier, C., Beno, N., Salles, C., & Thomas-Danguin, T. (2013). Enhancing salty taste through odour-taste-taste interactions: Influence of odour intensity and salty tastants' nature. *Food Quality and Preference*, *28*, 134–140.

Nguyen, H. D. (2000). Contribution l'étude des interactions entre les entres sensorielles: Effets d'une odeur sur la perception d'une saveur. [Ph.D. thesis, University of Burgundy].

Niimi, J., Eddy, A. I., Overington, A. R., Heenan, S. P., Silcock, P., Bremer, P. J., & Delahunty, C. M. (2014). Aroma-taste interactions between a model cheese aroma and five basic tastes in solution. *Food Quality and Preference*, *31*, 1–9.

Prescott, J. (1999). Flavour as a psychological construct: Implications for perceiving and measuring the sensory qualities of foods. *Food Quality and Preference*, *10*, 349–356.

Prescott, J., Johnstone, V., & Francis, J. (2004). Odor-taste interactions: Effects of attentional strategies during exposure. *Chemical Senses*, *29*, 331–340.

Prescott, J., & Murphy, S. (2009). Inhibition of evaluative and perceptual odour–taste learning by attention to the stimulus elements. *The Quarterly Journal of Experimental Psychology*, *62*, 2133–2140.

Proserpio, C., Laureati, M., Invitti, C., Cattaneo, C., & Pagliarini, E. (2017). BMI and gender related differences in cross-modal interaction and liking of sensory stimuli. *Food Quality and Preference*, *56*, 49–54.

Proserpio, C., Laureati, M., Invitti, C., Pasqualinotto, L., Bergamaschi, V., & Pagliarini, E. (2016). Cross-modal interactions for custard desserts differ in obese and normal weight Italian women. *Appetite*, *100*, 203–209.

Rousseau, B. (2004). Sensory evaluation techniques. In L. M. L. Nollet (Ed.), *Handbookf of food analysis: Physical characterization and nutrient analysis* (pp. 21–37). Marcel Dekker, Inc.

Spence, C. (2022). Factors affecting odour-induced taste enhancement. *Food Quality and Preference*, *96*, 104393.

Stevens, S. (1958). Adaptation-level vs. the relativity of judgment. *The American Journal of Psychology*, *71*, 633–646.

Stevenson, R. J., Boakes, R. A., & Prescott, J. (1998). Changes in odor sweetness resulting from implicit learning of a simultaneous odor-sweetness association: An example of learned synesthesia. *Learning and Motivation*, *29*, 113–132.

Stevenson, R. J., Prescott, J., & Boakes, R. A. (1995). The acquisition of taste properties by odors. *Learning and Motivation*, *26*, 433–455.

Strigler, F., Touraille, C., Sauvageaot, F., Barthélémy, J., Issanchou, S., & Pagès, J. (2009). Les épreuves discriminatives et descriptives. In F. Depledt (Ed.), *SSHA Evaluation Sensorielle Manuel Méthodologique* (pp. 99–200). Lavoisier.

Sinding, C., Saint–Eve, A., & Thomas-Danguin, T. (2022). Multimodal sensory interactions. In E. Guichard & C. Salles (Eds.), *Flavor from food to behaviors, wellbeing and health* (pp. 205–231). Woodhead Publishing Series in Food Science, Technology and Nutrition.

Wang, G., Bakke, A. J., Hayes, J. E., & Hopfer, H. (2019). Demonstrating cross-modal enhancement in a real food with a modified ABX test. *Food Quality and Preference*, *77*, 206–213.

Wang, G., Hayes, J. E., Hopfer, H., Ziegler, G. R., & Roberts, R. F. (2018). Dose-response relationships for vanilla flavor and sucrose in skim milk: Evidence of synergy. *Beverages.*, *4*, 13–21.

Welge-Lüssen, A., Drago, J., Wolfensberger, M., & Hummel, T. (2005). Gustatory stimulation influences the processing of intranasal stimuli. *Brain Research*, *1038*, 69–75.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Aveline, C., Thomas-Danguin, T., & Sinding, C. (2023). Ranking task versus intensity scales, what is the best method to evaluate odor-induced taste enhancement? *Journal of Sensory Studies*, e12857. https://doi.org/10.1111/joss.12857