Food Quality and Preference 112 (2023) 105027

Contents lists available at ScienceDirect

Food Quality and Preference

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

Comparing traditional check-all-that-apply (CATA) and implicit response time Go/No-go approaches for profiling consumer emotional response when tasting food

Maheeka Weerawarna N.R.P. [a], A. Jonathan R. Godfrey [b], Malcolm Loudon [c], Meika Foster [d], Joanne Hort [a,e,*]

[a] Food Experience and Sensory Testing (Feast) Lab, School of Food and Advanced Technology, Massey University, Palmerston North 4410, New Zealand
[b] Statistics Group, School of Mathematical and Computational Sciences, Massey University, Palmerston North 4410, New Zealand
[c] School of Psychology, Massey University, Palmerston North 4410, New Zealand
[d] Edible Research Limited, Ohoka 7475, New Zealand
[e] Riddet Institute, Massey University, Palmerston North 4410, New Zealand

ARTICLE INFO

Keywords:
Implicit response time
Emotional responses
Plant-based milk-alternatives
Cow's milk
Generalised linear mixed models
Product discrimination

ABSTRACT

Adapting implicit approaches to capture consumer responses when tasting foods is of recent interest in sensory and consumer science. Implicit consumer responses are reported to be more spontaneous than explicitly gathered data. Traditionally, emotional response to foods is captured using explicit methods like check-all-that-apply (CATA). The present study aimed to compare an implicit response time (IRT) Go/No-go approach with traditional CATA to profile consumer emotional responses. Participants (n = 104) were consumers of, or willing to consume, cow's milk and plant-based milk alternatives (PBMA). Emotional responses for two cow milk and five PBMA products were evaluated across two sessions using IRT and CATA. The cow milk products were replicated across the sessions to allow consistency of response across sessions to be evaluated for each method. Data were collected using a bespoke single page web application (JavaScript, ECMA 2015). Data consistency across sessions (Spearman correlation (ρ)), emotion selection frequency (ρ and generalised linear models) and product discrimination (linear mixed models and correspondence analysis) were compared across the IRT and CATA approaches. Results showed high data consistency from both IRT and CATA across the two sessions ($\rho > 0.89$). Emotion selection frequency was also comparable across IRT and CATA. Interestingly, CATA was differentiating more between cow's milk products and IRT within the PBMA space. However, further investigations showed that fewer participants provided different responses in CATA than when under time pressure in the IRT Go/No-go. Additional investigations on the performance of explicit versus implicit methods, or their combination, are required across different product matrices to identify the optimum approach to capture consumer product experience.

1. Introduction

Consumer emotional responses are captured as a part of profiling holistic product experiences (Ng & Hort, 2015; Nijman et al., 2019; Spinelli & Monteleone, 2018). Often, these emotional responses are measured using check-all-that-apply (CATA) lexicons where consumers are asked to explicitly indicate their decision/choice regarding the emotions evoked by tasting a particular product (Ng, Chaya, & Hort, 2013), under no time pressure.

Decision making involves a combination of two cognitive pathways. One is a conscious pathway, where decisions/choices are made based on slower analytical thinking, revealing explicit responses. The other is a faster unconscious pathway, which is automatic and implicit (Kahneman, 2002). As decisions/choices are a combination of both implicit and explicit pathways, product experience is not solely explicit (Cervellon, Dube, & Knauper, 2007). In fact, explicitly gathered consumer responses are reported to be influenced by self-presentations such as social desirability (Maass, Castelli, & Arcuri, 2000; Matukin & Ohme, 2017; Ohme

* Corresponding author at: Food Experience and Sensory Testing (Feast) Lab, School of Food and Advanced Technology (SF&AT) PN 452, Massey University, Private Bag 11222, Palmerston North 4410, New Zealand.
E-mail address: j.hort@massey.ac.nz (J. Hort).

https://doi.org/10.1016/j.foodqual.2023.105027
Received 15 May 2023; Received in revised form 16 October 2023; Accepted 17 October 2023
Available online 21 October 2023

& Boshoff, 2019) and subsequent inconsistencies between reported consumer responses on lifestyle and their actual experiences or behaviours. For example, several earlier studies have shown that some participants who self-identify as vegetarians also report that they consume meat, fish or other animal meat-based products (Gossard & York, 2003; Perry, McGuire, Neumark-Sztainer, & Story, 2001; Robinson-O'Brien, Perry, Wall, Story, & Neumark-Sztainer, 2009; Rothgerber, 2014; Worsley & Skrzypiec, 1998). Further, Berridge & Winkielman (2003) highlighted the potential of unconscious implicit emotions impacting on behavior (affective reaction) independently of consumers' conscious awareness. For example, preference ratings for Chinese ideographs were affected by unconsciously presented happy or angry faces (affective stimuli) (Winkielman, Zajonc, & Schwarz, 1997). Winkielman, Berridge, & Wilberger (2000), as cited by Berridge & Winkielman (2003), have shown subliminal exposure to happy faces causing assessors to pour and consume 50 % more fruit flavoured beverage than for neutral faces and vice versa for angry faces. Watson, Clark, & Tellegen (1988) showed increased willingness to pay and wanting to consume a fruit beverage upon subliminal exposure to happy faces than with angry faces when the assessors were thirsty. Therefore, explicitly captured emotional responses, specifically if products are less stimulating, could be influenced by lack of awareness regarding the emotions experienced during product evaluation. In fact, it is important to include implicit measures in consumer studies to explore implicit consumer response and to understand how these compare with explicit response. However, comparative studies between explicit and implicit emotional responses of consumers to tasted foods are scarce.

Implicit approaches are widely used in market research to capture consumer engagement (Ohme & Boshoff, 2019; Soulikhan, Norré, & Cheikhrouhou, 2021) and attitudes (Matukin & Ohme, 2017). In comparison, application of implicit approaches in consumer sensory research are rare. However, recent developments in interdisciplinary approaches, such as technologies like artificial intelligence or immersive environments (Beurteaux, 2023) and neuroscience techniques in marketing (Kemp, Nyambayo, Rogers, Sanderson, & Villarino, 2021; Niedziela & Ambroze, 2021), have created opportunities to adapt methods from cognitive science (De Houwer & Moors, 2010) to capture consumer implicit response (Cliceri, Spinelli, Dinnella, Prescott, & Monteleone, 2018; Kraus & Piqueras-Fiszman, 2016; Mai & Hoffmann, 2017; Mai, Hoffmann, Hoppert, Schwarz, & Rohm, 2015; Werle, Trendel, & Ardito, 2013; Woodward & Treat, 2015) to food. Among the few studies that have attempted to do so, implicit emotional responses evoked by evaluating foods tend to be captured primarily via two approaches, 1) physiological responses (de Wijk & Noldus, 2021; Lagast, Gellynck, Schouteten, De Herdt, & De Steur, 2017) such as facial expressions, eye movement (Kostyra, Wasiak-Zys, Rambuszek, & Waszkiewicz-Robak, 2016) or 2) neurological responses such as brain activity (Horska, Bercik, Krasnodebski, Matysik-Pejas, & Bakayova, 2016). Capturing implicit emotional responses using such approaches are often invasive and require specialised equipment and hence are undertaken at premium cost.

Implicit responses are also captured using behavioural measures such as accuracy and reaction time key press responses (De Houwer, & Moors, 2010). Reaction times to measures are often applied in psychology as an implicit measure denoting level of association of responses to stimuli (Fazio, 2001; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Greenwald & Banji, 1995; Hermans, Houwer, & Eelen, 1994; Shelton & Martin, 1992). Implicit reaction times have recently been adopted in consumer research to capture affective responses. For example, using an Affective Misattribution Procedure (Payne, Cheng, Govorun, & Stewart, 2005), Pierguidi et al. (2023) measured 'yes'/ 'no' reaction times to trustworthiness of face images after tasting basic taste solutions, as an implicit indication of affective responses related to the varying quality and intensities of the taste solutions. Porcherot et al. (2022) used a 2-choice procedure and profiled consumer emotional response to perfumes where participants were asked to respond as fast as possible to indicate whether the emotion appearing as a written word on screen was evoked by the perfume sample or not whilst software recorded the answers and respective reaction times. What is deemed an implicit response varies across studies, based on different methods employed. Porcherot et al. (2022) captured reaction times when researching fragrances as those under 3000 mS, longer than the 200 to 2500 mS reaction time adopted by Pellicer-Sánchez and Schmitt (2012) in a lexical decision-making task. Pierguidi et al. (2023) used a narrower 250 to 1500 mS range concurring with Hoogeveen, Jolij, ter Horst, and Lorist (2016). The Go/No-go method, another approach, measures response time to evaluate performance in paradigms where assessors either select the stimulus presented if relevant 'Go' or otherwise do not respond at all "No-go (Donders, 1969). Time elapsed from the appearance/occurrence of the stimuli until responding 'Go' is captured as an implicit measure denoting the congruency of the response to stimuli (Gomez, Ratcliff, & Perea, 2007; Gavazzi et al., 2023). To date, application of implicit response time Go/No-go tasks as a measure of consumer emotional response to tasting foods is not evident in the literature.

Comparative studies between traditional explicit approaches like CATA with implicit approaches are needed to investigate whether similar or different information is obtained from each identify the relative merits of each approach in profiling consumer emotional responses in different contexts. However, assessor performance and data formats are key considerations in such comparative studies. Within assessor reproducibility has been reported for sensory CATA (Jaeger & Ares, 2023; Jaeger et al., 2013) but no study has investigated this when using CATA to measure emotional responses. Assessor reproducibility concerning implicit approaches for measuring emotional responses is not conclusive. In terms of data formats, CATA produces binary data indicating whether an emotion was selected (=1) or not (=0) (Ares et al., 2017), whilst the implicit approach primarily utilise time to respond'Go' or 'No-go' to a stimuli, and is treated as continuous scale data (Gomez et al., 2007). In practice, the data formats used in CATA versus implicit response time are likely to make comparative studies challenging.

The current study aimed to evaluate the relative merits of using an Implicit Response Time (IRT) Go/No-go task and traditional CATA for profiling consumer emotional response when tasting using milk and plant-based milk alternatives (PBMA) as the food stimulus. More specifically, the key objectives were to compare the approaches for 1) assessor emotional response consistency across sessions, 2) emotion selection frequency, and 3) level of product discrimination, as well as to determine 4) whether implicit response time provides additional insights beyond binary CATA.

2. Methodology

2.1. Participants

This study was submitted to Massey University's Human Ethics process and was judged to be low risk (application ID 4000025636). A consumer panel was recruited from the Food Experience and Sensory Testing (Feast) laboratory consumer database, Massey University, New Zealand. Consumers (n = 105) aged 18 – 65 years, who were willing to consume cow's milk and PBMA, not pregnant or lactating and not allergic to the study products were selected to participate. A gift voucher was offered to participants to compensate for their time.

2.2. Products

This research was part of a wider project on understanding consumer response to alternatives to animal products where the shift in diet can evoke emotional response (Cardello, Llobell, Giacalone, Roigard, & Jaeger, 2022; Hielkema & Lund, 2021; Michel, Hartmann, & Siegrist, 2021; Michel, Knaapila, Hartmann, & Siegrist, 2021). A comparison between cow milk and its alternatives was hypothesised to show

Table 1
Products evaluated in the study.

Product	*Product description as shown on packaging
CowHF	Anchor™ blue (full cream)
CowLF	Anchor™ light (low-fat)
Oat	Sanitarium™ So Good Oat No Added Sugar
Soy	Alpro® Barista for Professionals Soy
Rice	Vitasoy® Rice Milk Unsweetened
Cashew	Sanitarium™ So Good Cashew Unsweetened
Almond	Sanitarium™ So Good Almond Unsweetened

*All products were ultra-high temperature (UHT) processed, unsweetened and unflavoured.

Table 2
Emotion/cognitive product conceptualisation lexicon and example statements given to consumers.

Emotion Term		The beverage made me feel…
1	Adventurous	Daring, brave, seeking, courageous
2	Bored	Uninterested, fed up
3	Cheap	Inferior, unworthy
4	Classy	Elegant, stylish,
5	Comforted	Consoled, reassured, pacified
6	Energised	Eager, rejuvenated, invigorated
7	Feminine	Womanly, womanlike
8	Genuine	Honest, sincere
9	Happy	Content, joyful
10	Inspired	Enthused, stimulated
11	Irritated	Annoyed, aggravated
12	Masculine	Manly, macho
13	Modern	Contemporary, fashionable
14	Pretentious	Showy, hollow, fake
15	Relaxed	Tranquil, calm
16	Sensual	All my senses are excited
17	Simple	Uncomplicated, not overly fussy
18	Sophisticated	Superior, advanced, mature
19	Traditional	Conventional, customary
20	Uninspired	Dull, lacklustre

response differences. Further as New Zealand is a prominent dairy manufacturer and milk forms a primary component in the diet (Stats-NZ, 2022), such research is very relevant to the New Zealand economy where the research was funded. Therefore, two cow milks and five PBMA commercially available in New Zealand (Table 1) were evaluated.

2.3. Emotion measurement

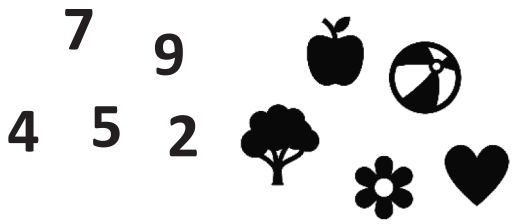
The emotion and cognitive product conceptualisation lexicon used by Cardello, Llobell, Giacalone, Roigard, and Jaeger (2022) for assessment of PBMA and cow milk with New Zealand consumers was adopted for this study. Supplementary example statements for each emotion term (Table 2) were developed using Microsoft 365 thesaurus (version 2212) and bench testing session with Feast team members. These were used during the session introduction to aid consumer interpretation of emotion terms.

Data were collected using iPads with a keyboard attached via a lightening connector using a bespoke single page web application (JavaScript, ECMA 2015) developed at Massey University.

2.3.1. CATA approach

Participants tasted a 10 mL sip of product and selected all emotion terms (Table 2) from an on-screen CATA list that they felt the sample made them feel. Emotion term order was randomised across participants according to a balanced William Latin Square design to avoid attribute position and order related confounding effects and was presented in five columns. However, term order was fixed within a given participant for all the product evaluations to avoid frustration associated with terms moving around and enable higher operational power (Meyners & Castura, 2016).

Table 3
Graphics and matching words used in Step 2 – False positive rate testing in the IRT Go/No-go approach.

Graphics	Numbers	Objects
		
Matching words	*Seven, Nine, Five, Two, Four*	*Tree, Flower, Ball, Heart, Apple*

2.3.2. IRT Go/No-go approach

The IRT Go/No-go approach comprised three steps: Introduction, false positive rate testing and product evaluation. Introduction and false positive rate testing were performed once only at the beginning of each product evaluation session.

Step 1 – Introduction of Go/No-go task: The Go/No-go concept was introduced to participants here. Participants were instructed to press the spacebar as soon as possible every time they saw a star symbol (★) appear on the screen. The symbol appeared ten times with random time gaps fixed across participants but varying from 400 mS to 1000 mS. Additionally, a 500 mS break occurred after the space bar was pressed and before the next appearance of the symbol following Gomez, Ratcliff, and Perea (2007). At each appearance, the symbol remained on the screen for 1000 mS. This 1000 mS cut-off time was selected to represent the maximum time regarded as an implicit response to the onscreen stimulus throughout this study. This is longer than the minimum of 200 mS (Thompson et al., 1992) to 233 mS (Krieg Jr et al., 2001) simple reaction time (i.e., where assessors respond as fast as possible to the occurrence of a stimulus) and less than the maximum cut-off time of 3000 mS used by Porcherot et al. (2022) but was deemed a reasonable high cut-off in pilot testing. If the space bar was not pressed within 1000 mS an error message appeared on the screen instructing the participant to press the space bar as soon as possible and continue.

Step 2 – False positive rate testing: This step was used to evaluate whether participants were paying attention to the task and being cautious or lax (Kantner & Lindsay, 2012; Wixted, 2020) concerning when they press the space bar to avoid mistakes or missing a correct response, and to provide a practice of a Go/No-go approach. A graphic of either a collection of numbers or objects appeared on the screen with one of ten words related to those numbers or objects (Table 3).

Participants were instructed to press the space bar (Go) if the word matched the graphic on screen and do nothing (No-go) if not and wait for the next word to appear. The words appeared ten times per graphic (five out of ten were correct matches to the graphic) followed by a second set with the other graphic. Word presentation order was balanced across participants.

Step 3 – Product evaluation: Participants were instructed to taste a 10 mL sip of a product and press the space bar to begin. Once pressed, emotion terms (Table 2) appeared monadically on the screen. Participants were instructed to press the space bar (Go) if the sample they tasted made them feel the onscreen and do nothing (No-go) and wait for the next term if it was not relevant. A 500 mS time gap existed between pressing the space bar and the next emotion word appearing (Gomez et al., 2007). If a participant did not press the space bar within 1000 mS the term remained on screen and automatically forwarded to the next emotion term. The time to press the space bar was recorded from the initial appearance of the term. A time of 1000 mS was recorded to represent 'No-go' when a participant failed to press the space bar by 1000 mS.

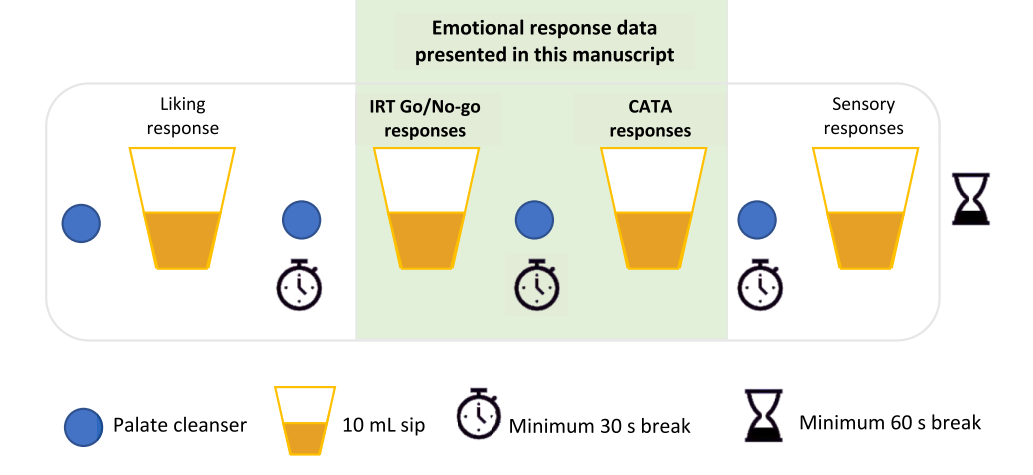


Fig. 1. Sample evaluation protocol per product. The order of CATA and IRT Go/No-go emotional response tasks was balanced between participants to remove an order effect.

2.4. Evaluation sessions

Product evaluations were carried out in the Feast lab at Massey University, New Zealand. Samples ($5 \pm 1\ ^0$C) were evaluated in standard sensory booths (ISO8589:2007(E)) under white lighting at a room temperature of $21 \pm 1\ ^0$C.

Participants attended two 1.5 h sessions. The two cow milk products were repeated across both sessions to allow consistency of emotional response to be investigated. It was hypothesised if inconsistent data were noticed across sessions, a familiar product would better represent the session effect than repeated exposure effects (Anzman-Frasca & Ehrenberg, 2018; Tempere et al., 2019). Given cow milk is a familiar product for NZ consumers with respect to plant-based beverages, cow milk samples were replicated across sessions to investigate if a session effect existed. The remaining five samples were split and evaluated across the two sessions, meaning participants assessed four or five samples in a session. Products were served according to a randomised balanced design across the sessions.

Participants evaluated four 10 mL sips per product (Fig. 1). The first and fourth sip were for data collection on liking and sensory characteristics as part of a wider research programme. In between they assessed two 10 mL samples to indicate their emotional response, the data relevant to this study. Each 10 mL sip was presented in a separate 30 mL cup labelled with a random three-digit code. To minimise carryover effects, participants were instructed to palate cleanse with filtered water and a bite of cracker (Arnott's water cracker original, Australia) prior to tasting from a cup and to take a forced minimum break of 30 s after each 10 mL sip evaluation. A forced minimum break of 60 s was imposed prior to evaluation of each subsequent product. Half of the participants completed the CATA task first and the other half completed the IRT Go/No-go task first.

2.5. Data analysis

Statistical analyses were performed using R software, version 4.2.2 (R Core Team, 2022) in RStudio (2022.12.0) with $\alpha = 0.05$. Package `dplyr` (Wickham, François, Henry, & Müller, 2020) was used for data handling and `ggplot2` for data visualisation. Specific R-packages used for analyses are detailed under relevant sections below.

2.5.1. Data pre-processing

CATA: Mean citation proportions and observed standard error (SE) for all CATA emotion terms were calculated for each product.

IRT: Data from IRT Go/No-go step 2 – false positive rate testing was used to identify if any participants should be removed from the data set based on false answers given to all the answering attempts.

Step 3 – product evaluation in IRT Go/No-go task creates two types of data; 1) response time (implicit measure) = time taken to press the space bar for evoked emotions (continuous data) and 2) Go/No-go (explicit measure) = indicates whether the space bar was pressed or not for respective emotion terms (binary data). Both the response time data and the Go/No-go data formats are analysed here.

To standardise the response time data to account for natural variation in individual response time (Ratcliff, 1993), a baseline for individual response time (iRT) was identified by extracting the minimum response time from overall evaluations and was used in further analysis to calculate standardised RT (Std.IRT) for each participant as shown in Equation (1).

$$\text{Std.IRT} = \frac{(Implicitresponsetime - minRT)}{(1000 - minRT)} \quad (1)$$

Mean Std.IRT and SE were calculated for each emotion term per product for further analysis.

2.5.2. Consistency of emotional response across sessions for each approach

To ascertain if emotion responses were consistent across sessions, Spearman correlation coefficients (ρ) were calculated between the mean citation proportions from CATA data collected in sessions 1 and 2 for the two cow milk samples. Similarly, ρ was calculated between the mean Std.IRT of emotion terms from session 1 and 2. To visualise the data, responses from session 1 were plotted against session 2 for CATA and IRT data.

2.5.3. Frequency of emotion selection across approaches

Go/No-go data was identified as Go (=1) or No-go (=0) for each emotion and product. Mean emotion citation proportions of Go/No-go were calculated and plotted against mean citation proportions of the CATA data and ρ was calculated between the frequencies of the CATA and Go/No-go.

Then, the ρ was also calculated between CATA citation proportions versus Std.IRT across all emotions to determine if response time provided additional information than CATA data. To achieve this, mean citation proportions for CATA was plotted against mean Std.IRT to visualise the data. Further, investigations were made by plotting mean CATA citation proportions versus mean Std.IRT per emotion.

To determine consistency of emotion selection, cross tabulations were created to obtain and compare counts and proportions of CATA (0 = not-checked, 1 = checked) responses and Go/No-go data (0 = No-go,

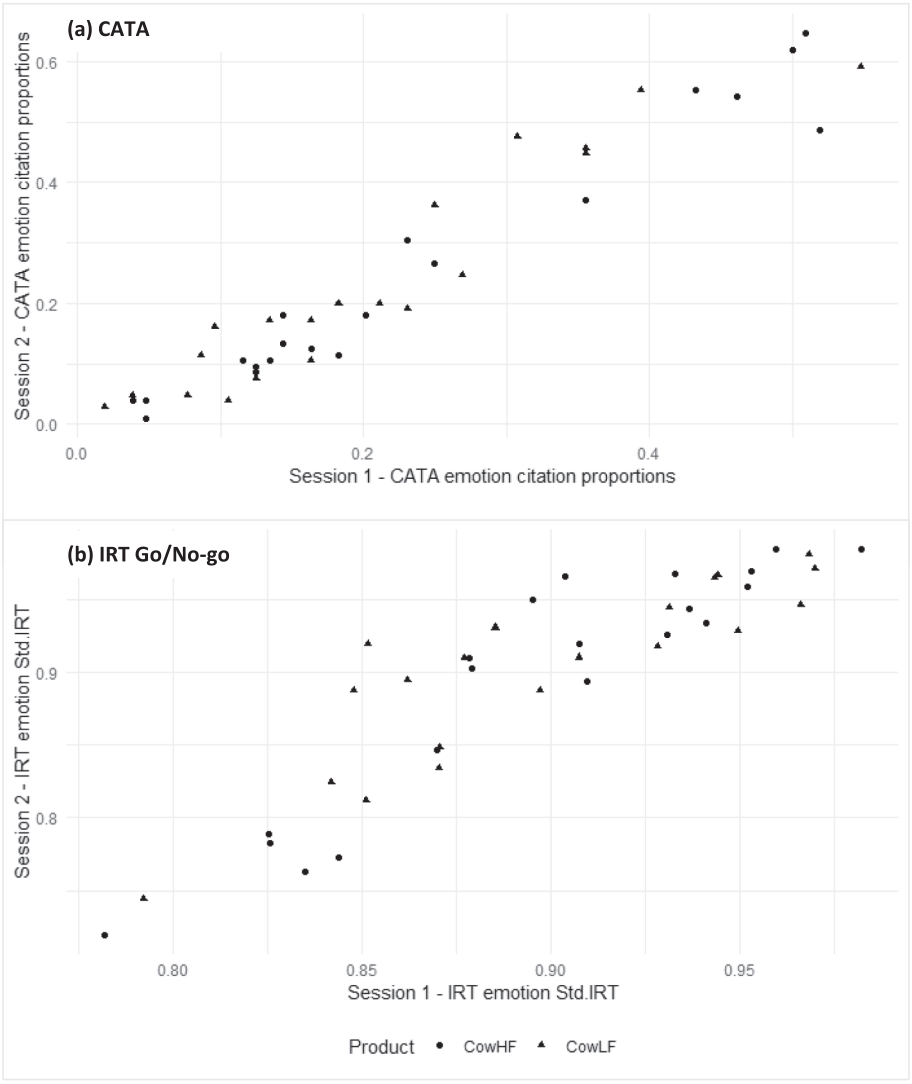


Fig. 2. Scatterplots of emotion terms (individual points represent individual emotions) from session 1 and 2 of the two cow milk products (shown by the circle (COWHF) and triangle (COWLF) points) (a) CATA mean citation proportions, $\rho = 0.962$, (b) IRT Go/No-go mean standardised implicit response time (Std.IRT), $\rho = 0.897$.

1 = Go) for pooled data, by individual emotions and products. A two-factor (*emotion, inconsistency*) generalised linear model (GLM) and Analysis of Deviance (Agresti, 2018; Weerawarna N. R. P., Godfrey, Ellis, & Hort, 2021a) was performed on the inconsistent proportions of No-go = 0 when CATA = 1, and Go = 1 when CATA = 0, i.e. a subset of the data. A second two-factor (*product, inconsistency*) GLM and Analysis of Deviance was performed on the same subset of the data to test the product effect. Tukey HSD post-hoc tests were used to make pairwise comparisons of the proportions where main effects were significant.

2.5.4. Product discrimination

Two approaches, 1) linear mixed models (Package `lme4` (Bates et al., 2015)), and 2) Correspondence Analysis or Principal Component Analysis (Package `FactoMineR` (Le et al., 2008)) were used to analyse the level of product discrimination in CATA versus IRT approaches.

2.5.4.1. Linear mixed models. First, linear mixed models and subsequent post-hoc mean comparisons were used to investigate the product effects on emotional responses.

CATA: GLM with binomial distributions are recommended to analyse binary data (Agresti, 2018; Bi & Kuesten, 2022; Weerawarna N. R. P. et al., 2021a). Mixed logistic regression was performed through Generalised linear mixed effect models (GLME) with default functions (function = glmer, family="binomial", link = Logit; (Agresti, 2018) were used to model the citation proportions for *product* as fixed effect and *participant* as random effect to investigate *product* effects on emotional responses. Analysis of Deviance (Agresti, 2018) was used to determine the statistically significance of *product*. Tukey HSD post-hoc test was used to make pairwise comparisons of citation proportions where emotional responses were significant among *products*.

IRT: Response time is inversely proportional to the strength of association of an emotion to a product (Gomez, Ratcliff, & Perea, 2007; Gavazzi et al., 2023). Therefore, Inverted Std.IRT (iStd.IRT) = 1 – Std. IRT of an emotion term, was used in this analysis. Mean iStd.IRT and observed SE for all emotion terms were calculated for each product. Typically, continuous measures of response time data from IRT Go/No-go method are analysed using ANOVA (Gomez et al., 2007). Linear mixed effect models (LMM) were used to model the iStd.IRT for *product* as fixed effect and *participant* as random effect to investigate product effects on emotional responses. ANOVA was used to determine the statistical significance of *product*. Where emotional responses were significant among *products*, Tukey HSD post-hoc pairwise comparisons of mean iStd.IRT were applied.

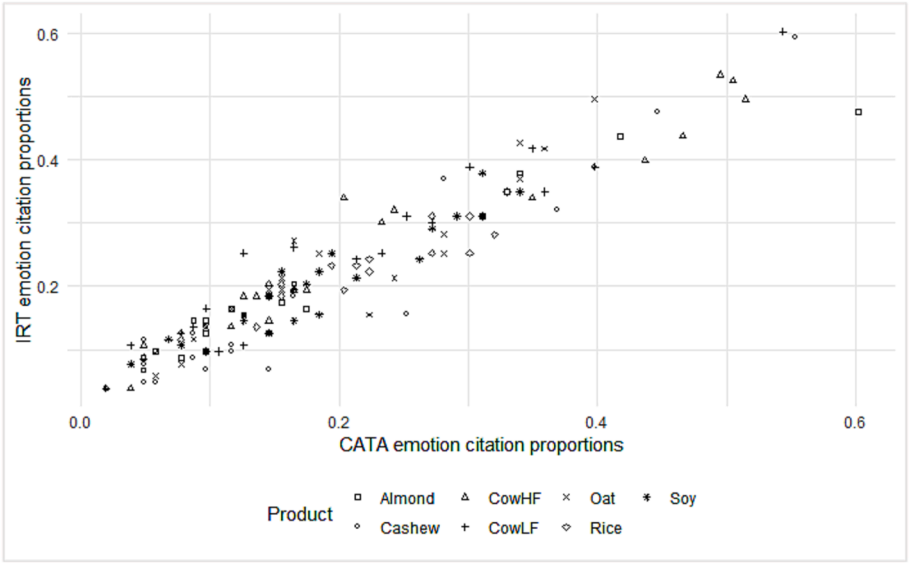


Fig. 3. Scatterplot of the mean citation proportions of emotion terms (individual points represent individual emotions) in CATA versus IRT Go/No-go approaches of products (shown in different shape points). $\rho = 0.947$.

2.5.4.2. Correspondence Analysis and Principal Component Analysis. **CATA:** Secondly, Correspondence Analysis and subsequent biplots were obtained to visualise the CATA emotional response space of the products. A product × emotion terms contingency matrix was developed (Meyners, Castura, & Carr, 2013) for CATA citation proportions on which Correspondence Analysis (Abdi & Béra, 2014) was performed.

IRT: As longer response times are associated with lower relevance of an emotion to a product (Gomez, Ratcliff, & Perea, 2007; Gavazzi et al., 2023), a product × emotion contingency matrix was developed using iStd.IRT data on which Principal Component Analysis was performed.

Then, subsequent biplots were created for each CATA and IRT. The chi-square distance was obtained to map the biplot of product × emotion terms. Dimensions were retained that enabled at least 80 % of the variance in the data to be retained. Positioning of product confidence ellipses (95 %) were viewed to determine if emotional response profiles of products were differentiated.

Euclidean distance matrices of the biplots from CATA and IRT data were obtained. Pearson's correlation was used to investigate the relationship between the distance matrices and hence level of product discrimination in CATA and IRT.

3. Results

The false positive rate testing task in the IRT Go/No-go method indicated one participant gave all false answers (answering 'Go' for 'No-go' and vice versa) indicating lack of attention or misunderstanding the task. Their data were removed from the analysis and hence the final analysis was based on 104 participants.

3.1. Consistency of emotional response across sessions for each approach

Mean CATA citation proportions and mean Std.IRT from session 1, showed high correlations ($\rho > 0.89$) with their respective session 2 data (Fig. 2 (a & b)). No session effect on emotional responses in both CATA

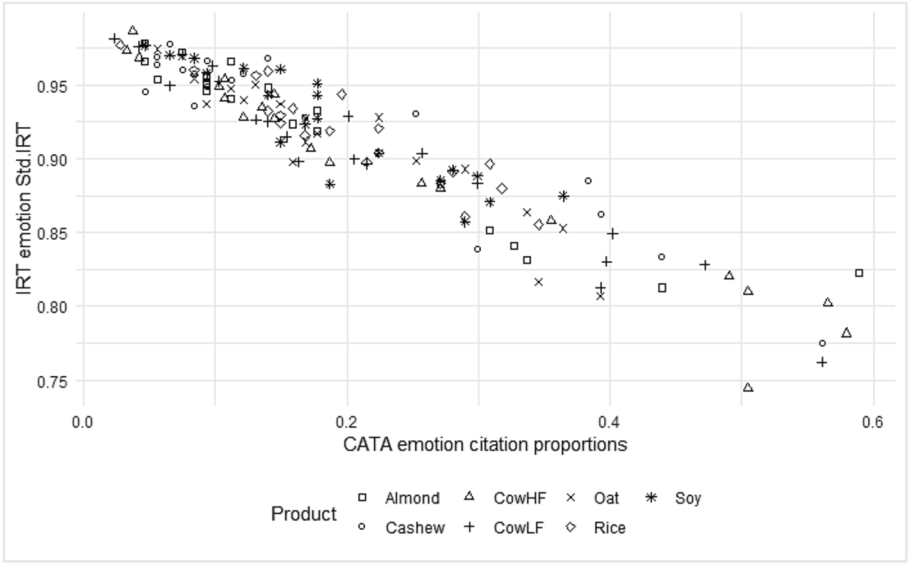


Fig. 4. Scatterplot of mean CATA citation proportions and mean Std.IRT of IRT Go/No-go (products shown in different shape points) across all emotion terms (individual points represent individual emotions), $\rho = -0.949$.

Fig. 5. Scatterplot of mean CATA citation proportions versus mean Std.IRT of IRT Go/No-go per emotion term across products (shown in different shape points). Redundant emotion terms highlighted by circles.

Table 4
Cross tabulation of CATA and Go/No-go proportions (pooled data).

		Go/No-go	
		0 (No-go)	*1 (Go)*
CATA	*0 (not checked)*	0.727	0.051
	1 (checked)	0.075	0.146

Shaded cells show inconsistent responses across the two methods.

and IRT data was evident.

3.2. Frequency of emotion selection across approaches

Citation proportions across all emotions in CATA and Go/No-go (Fig. 3) were highly correlated ($\rho = 0.947$).

Citation proportions from CATA and Std.IRT data displayed a high negative correlation (ρ = -0.949) (Fig. 4) indicating lower citation proportions of emotion terms in CATA related to longer response times. Further investigations at the individual emotion level (Fig. 5) revealed the potential redundancy of some terms. For example, 'masculine', 'sensual' and 'sophisticated' had very low citation proportions in CATA (<0.2) and if they were also chosen in IRT they had long response times (>900 mS) across all products.

Table 4 cross-tabulates CATA and Go/No-go response proportions on pooled data indicating that, 88 % of the time, participant responses were consistent across methods i.e. 'unchecked' in CATA, and 'No-go'; or 'checked' in CATA and 'Go'. However, although limited, there were some inconsistent participants a proportion of 'No-go' responses corresponding to being 'checked' in CATA (0.075) and that of 'Go' and CATA 'not-checked' (0.051). Table 5 and Table 6 detail the cross tabulation by

Table 5
Cross tabulation of CATA and Go/No-go proportions by product.

Product	Go/No-go	CATA	
		0	*1*
CowHF	*0*	0.68	0.09^{A}
	1	0.05^{B}	0.18
CowLF	*0*	0.7	0.09^{A}
	1	0.05^{B}	0.15
Almond	*0*	0.75	0.07^{A}
	1	0.05^{B}	0.13
Cashew	*0*	0.77	0.05^{A}
	1	0.04^{B}	0.13
Oat	*0*	0.71	0.09^{A}
	1	0.06^{B}	0.14
Rice	*0*	0.73	0.07^{A}
	1	0.05^{B}	0.15
Soy	*0*	0.74	0.07^{A}
	1	0.05^{B}	0.14

[AB] 'inconsistent' ('No-go'=0 and 'checked'=1 in CATA or 'Go'=1 and 'not-checked'=0 in CATA) diagonal proportions (shaded cells) with different letter codes within a product are significantly different (*Tukey HSD, $p < 0.05$*).

Table 6
Cross tabulation of CATA and Go/No-go proportions by emotion term.

Emotion term	Go/No-go	CATA	
		0	*1*
Adventurous	*0*	0.78	0.06^{A}
	1	0.04^{A}	0.13
Bored	*0*	0.68	0.09^{A}
	1	0.06^{B}	0.17
Cheap	*0*	0.68	0.07^{A}
	1	0.06^{A}	0.19
Classy	*0*	0.82	0.07^{A}
	1	0.03^{B}	0.08
Comforted	*0*	0.66	0.08^{A}
	1	0.06^{A}	0.2
Energised	*0*	0.77	0.07^{A}
	1	0.03^{B}	0.13
Feminine	*0*	0.85	0.05^{A}
	1	0.03^{A}	0.07
Genuine	*0*	0.72	0.09^{A}
	1	0.07^{A}	0.12
Happy	*0*	0.65	0.08^{A}
	1	0.04^{B}	0.22
Inspired	*0*	0.75	0.11^{A}
	1	0.03^{B}	0.11
Irritated	*0*	0.69	0.07^{A}
	1	0.04^{B}	0.2
Masculine	*0*	0.93	0.02^{A}
	1	0.01^{A}	0.03
Modern	*0*	0.67	0.11^{A}
	1	0.06^{B}	0.16
Pretentious	*0*	0.8	0.05^{A}
	1	0.06^{A}	0.09
Relaxed	*0*	0.64	0.09^{A}
	1	0.07^{A}	0.2
Sensual	*0*	0.84	0.05^{A}
	1	0.03^{A}	0.08
Simple	*0*	0.47	0.12^{A}
	1	0.12^{A}	0.29
Sophisticated	*0*	0.83	0.06^{A}
	1	0.04^{A}	0.06
Traditional	*0*	0.71	0.08^{A}
	1	0.07^{A}	0.13
Uninspired	*0*	0.59	0.06^{A}
	1	0.08^{A}	0.27

^{AB}'inconsistent' ('No-go'=0 and 'checked'=1 in CATA or 'Go'=1 and 'not-checked'=0 in CATA) diagonal proportions (shaded cells) with different letter codes within an emotion are significantly different (*Tukey HSD, $p < 0.05$*).

product and emotion term respectively. Analyses of Deviance showed significant effects of *inconsistent* (Pr(>Chi) < 0.010), *product* (Pr(>Chi) < 0.019) and *emotion term* (Pr(>Chi) < 0.001) on the proportions. The *product* effect was consistent across products (Table 5). However, significant interaction of *inconsistency* and *emotion terms* (Pr(>Chi) < 0.001) indicated that inconsistencies are emotion dependent. Mean comparisons showed a significant size effect across inconsistent proportions for 'bored', 'classy', 'comforted', 'energised', 'happy', 'inspired', 'irritated' and 'modern' (Table 6).

3.3. Product discrimination

Table 7 shows the mean citation proportions and SE of CATA emotion terms by product. Table 8 shows mean iStd.IRT and SE of emotion terms by product.

Analysis of Deviance on GLME of CATA revealed significant *product* effects on emotion terms except for 'classy' and 'masculine' (Table 9), although 'masculine' was approaching significance (Pr(>Chi) = 0.098). In comparison, ANOVA on LMM of mean iStd.IRT revealed a significant *product* effect for emotion terms, except 'feminine', 'masculine', 'modern', 'pretentious' and 'sophisticated' (Table 9), however 'feminine' (Pr(>F) = 0.075) and 'modern' (Pr(>F) = 0.092) were approaching significance. In fact, number of emotion terms with significant *product* effect in IRT is less than in CATA. Additionally, *product* effect on 'classy' is not significant in CATA but in IRT. When considering the pair-wise mean comparisons on significant terms, within cow milks, CATA differentiated CowHF from CowLF on more emotion terms ('adventurous', 'comforted', 'inspired', 'irritated', 'relaxed' and 'uninspired') where IRT only differentiated on 'simple'. However, vice versa was noticed within PBMA space. IRT was differentiating products slightly more on 'bored', 'classy', 'comforted', 'happy' and 'simple' than CATA. For example, lower CATA citations for 'comforted' of Almond and Cashew only differed from higher citations of 'Oat'. In IRT, lower mean score for 'comforted' of 'Soy' was additionally differentiated from 'Oat'. However, with an exception, CATA differentiated 'energised' across PBMA marginally more, i.e. lower CATA citations of Cashew from Oat, Rice and Soy, but not in IRT. Both CATA and IRT did not differentiate the means for 'sensual' across the products.

Correspondence Analysis biplots for CATA (Fig. 6 (a)) and Principal Component Analysis for IRT (Fig. 6 (b)) revealed similar emotional response spaces for products and a high correlation between the distance matrices from CATA and IRT (Pearson's correlation coefficient = 0.986), indicated a similar degree of product differentiation from CATA and IRT. However, the 95 % confidence ellipses for the iStd.IRT data overlapped more than observed for the CATA data (Fig. 6), suggesting CATA amplifies product discrimination.

4. Discussion

This study explored whether an IRT Go/No-go approach provides similar or different insights into consumer emotional responses with respect to a traditional explicit CATA approach.

Within assessor reproducibility and stable product characterisation have been reported for sensory CATA (Jaeger et al., 2013). This study also shows assessor reproducibility when using CATA for emotional response measurements. For the first time, consistency of participant IRT data across sessions was identified in this study. It is important to note that assessor consistency across sessions was evaluated based only on the replicates of the two cow milk samples. Overall IRT and CATA performed similarly well in terms of assessor consistency, indicating that testing number of products across several sessions for emotional response is appropriate, and possible regardless of the choice of CATA versus IRT Go/No-go.

Selected emotions, and frequency of their selection, were also similar across the two approaches, indicating that the choice of CATA versus IRT Go/No-go had no methodological impact on recording emotions. Moreover, correlation between the continuous response time measure with citation proportions in CATA supports the assumption that faster reaction times indicate more congruent relationships between the stimuli as observed by Verhulst et al. (2006). Similar findings were reported by others investigating the association of implicit reaction time with elicited emotions for perfumes (Porcherot et al., 2022) and congruent relationship between the chemosensory quality and intensity of sensations and implicit affective reactions (Pierguidi et al., 2023).

Further, in the present study, the high negative correlation of response time and citation proportions of CATA implies that where

Table 7
Mean CATA citation proportions and standard error (*SE*) of emotion terms by product.

Emotion terms	Products													
	CowHF		CowLF		Almond		Cashew		Oat		Rice		Soy	
	Mean citation proportions and *SE*													
	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*
Adventurous	0.15^{b}	*0.02*	0.06^{a}	*0.03*	0.10^{ab}	*0.03*	0.11^{ab}	*0.03*	0.29^{c}	*0.04*	0.22^{bc}	*0.04*	0.22^{bc}	*0.04*
Bored	0.12^{a}	*0.04*	0.21^{ab}	*0.03*	0.35^{c}	*0.05*	0.30^{bc}	*0.04*	0.16^{ab}	*0.04*	0.16^{ab}	*0.04*	0.26^{bc}	*0.04*
Cheap	0.11^{a}	*0.04*	0.20^{ab}	*0.03*	0.33^{bc}	*0.05*	0.41^{c}	*0.05*	0.12^{a}	*0.03*	0.28^{bc}	*0.04*	0.28^{bc}	*0.04*
Classy	**0.16^{a}**	***0.03***	**0.13^{a}**	***0.04***	**0.10^{a}**	***0.03***	**0.10^{a}**	***0.03***	**0.11^{a}**	***0.03***	**0.15^{a}**	***0.04***	**0.07^{a}**	***0.02***
Comforted	0.58^{d}	*0.05*	0.40^{c}	*0.05*	0.09^{a}	*0.03*	0.11^{a}	*0.03*	0.33^{bc}	*0.05*	0.21^{ab}	*0.04*	0.18^{ab}	*0.04*
Energised	0.27^{c}	*0.04*	0.20^{bc}	*0.04*	0.10^{ab}	*0.03*	0.05^{a}	*0.02*	0.18^{bc}	*0.04*	0.19^{bc}	*0.04*	0.17^{bc}	*0.04*
Feminine	0.14^{ab}	*0.03*	0.10^{ab}	*0.03*	0.07^{ab}	*0.02*	0.05^{a}	*0.02*	0.09^{ab}	*0.03*	0.16^{b}	*0.04*	0.08^{ab}	*0.03*
Genuine	0.36^{c}	*0.05*	0.31^{c}	*0.05*	0.11^{ab}	*0.03*	0.06^{a}	*0.02*	0.25^{bc}	*0.04*	0.14^{ab}	*0.03*	0.14^{ab}	*0.03*
Happy	0.50^{d}	*0.05*	0.40^{cd}	*0.05*	0.09^{a}	*0.03*	0.09^{a}	*0.03*	0.34^{bc}	*0.05*	0.30^{bc}	*0.04*	0.19^{ab}	*0.04*
Inspired	0.26^{b}	*0.04*	0.15^{a}	*0.04*	0.09^{a}	*0.03*	0.05^{a}	*0.02*	0.16^{ab}	*0.04*	0.15^{ab}	*0.04*	0.15^{ab}	*0.04*
Irritated	0.04^{a}	*0.03*	0.13^{b}	*0.02*	0.30^{c}	*0.05*	0.46^{c}	*0.05*	0.10^{ab}	*0.03*	0.34^{c}	*0.05*	0.30^{c}	*0.04*
Masculine	**0.04^{a}**	***0.01***	**0.02^{a}**	***0.02***	**0.06^{a}**	***0.02***	**0.10^{a}**	***0.03***	**0.06^{a}**	***0.02***	**0.03^{a}**	***0.02***	**0.04^{a}**	***0.02***
Modern	0.19^{ab}	*0.04*	0.17^{a}	*0.04*	0.17^{ab}	*0.04*	0.16^{a}	*0.04*	0.28^{ab}	*0.04*	0.27^{ab}	*0.04*	0.31^{b}	*0.05*
Pretentious	0.03^{a}	*0.03*	0.07^{ab}	*0.02*	0.16^{bcd}	*0.04*	0.25^{d}	*0.04*	0.08^{ac}	*0.03*	0.20^{cd}	*0.04*	0.17^{cd}	*0.04*
Relaxed	0.56^{d}	*0.05*	0.40^{c}	*0.05*	0.15^{a}	*0.04*	0.09^{a}	*0.03*	0.36^{bc}	*0.05*	0.22^{ab}	*0.04*	0.18^{a}	*0.04*
Sensual	0.11^{a}	*0.03*	0.10^{a}	*0.03*	0.05^{a}	*0.02*	0.06^{a}	*0.02*	0.14^{a}	*0.03*	0.14^{a}	*0.03*	0.12^{a}	*0.03*
Simple	0.50^{bc}	*0.05*	0.57^{c}	*0.05*	0.41^{ab}	*0.05*	0.36^{ab}	*0.05*	0.40^{ab}	*0.05*	0.31^{a}	*0.05*	0.30^{a}	*0.05*
Sophisticated	0.10^{ab}	*0.02*	0.04^{a}	*0.03*	0.08^{ab}	*0.03*	0.14^{b}	*0.03*	0.12^{ab}	*0.03*	0.08^{ab}	*0.03*	0.14^{b}	*0.03*
Traditional	0.50^{b}	*0.05*	0.47^{b}	*0.05*	0.13^{a}	*0.03*	0.08^{a}	*0.03*	0.15^{a}	*0.04*	0.13^{a}	*0.03*	0.10^{a}	*0.03*
Uninspired	0.14^{a}	*0.04*	0.26^{b}	*0.03*	0.60^{c}	*0.05*	0.56^{c}	*0.05*	0.23^{ab}	*0.04*	0.30^{b}	*0.05*	0.35^{b}	*0.05*

^{abcd}Product citation proportions with different letter codes by row (emotion term) are significantly different (*Tukey, $p < 0.05$*).

Table 8
Mean inverted standerdised response time (iStd.IRT) (mS) and *standard error* (*SE*) (mS) of emotion terms by product.

Emotion terms	Products													
	CowHF		CowLF		Almond		Cashew		Oat		Rice		Soy	
	Mean iStd.IRT (mS) and *SE* (mS)													
	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*
Adventurous	0.09^{ab}	0.02	0.06^{ab}	0.02	0.05^{a}	0.01	0.04^{a}	0.01	0.11^{b}	0.02	0.08^{ab}	0.02	0.09^{ab}	0.02
Bored	0.09^{ab}	0.02	0.12^{bc}	0.02	0.17^{c}	0.02	0.16^{c}	0.02	0.06^{a}	0.02	0.08^{a}	0.02	0.11^{ab}	0.02
Cheap	0.05^{a}	0.01	0.12^{ab}	0.02	0.16^{b}	0.02	0.14^{b}	0.02	0.06^{a}	0.01	0.11^{ab}	0.02	0.11^{ab}	0.02
Classy	0.08^{b}	0.01	0.06^{ab}	0.01	0.03^{a}	0.01	0.05^{ab}	0.01	0.05^{ab}	0.01	0.08^{b}	0.01	0.03^{a}	0.01
Comforted	0.21^{c}	0.02	0.17^{bc}	0.02	0.05^{a}	0.01	0.05^{a}	0.01	0.14^{b}	0.02	0.10^{ab}	0.02	0.06^{a}	0.01
Energised	0.11^{b}	0.02	0.08^{ab}	0.02	0.05^{a}	0.01	0.03^{a}	0.01	0.08^{ab}	0.01	0.08^{ab}	0.02	0.08^{ab}	0.02
Feminine	**0.06^{a}**	**0.01**	**0.05^{a}**	**0.01**	**0.05^{a}**	**0.01**	**0.02^{a}**	**0.01**	**0.05^{a}**	**0.01**	**0.07^{a}**	**0.01**	**0.03^{a}**	**0.01**
Genuine	0.13^{c}	0.02	0.11^{bc}	0.02	0.06^{ab}	0.01	0.03^{a}	0.01	0.10^{bc}	0.02	0.07^{ab}	0.01	0.06^{ab}	0.01
Happy	0.23^{c}	0.02	0.17^{bc}	0.02	0.05^{a}	0.01	0.06^{a}	0.02	0.18^{bc}	0.02	0.14^{bc}	0.02	0.12^{bc}	0.02
Inspired	0.12^{b}	0.02	0.09^{ab}	0.02	0.05^{a}	0.01	0.06^{a}	0.01	0.10^{ab}	0.02	0.07^{ab}	0.01	0.09^{ab}	0.02
Irritated	0.03^{a}	0.01	0.09^{ab}	0.02	0.15^{bc}	0.02	0.17^{c}	0.02	0.06^{a}	0.02	0.14^{bc}	0.02	0.14^{bc}	0.02
Masculine	**0.01^{a}**	**0.01**	**0.03^{a}**	**0.01**	**0.03^{a}**	**0.01**	**0.03^{a}**	**0.01**	**0.03^{a}**	**0.01**	**0.02^{a}**	**0.01**	**0.02^{a}**	**0.01**
Modern	**0.11^{a}**	**0.02**	**0.11^{a}**	**0.02**	**0.08^{a}**	**0.02**	**0.07^{a}**	**0.02**	**0.12^{a}**	**0.02**	**0.12^{a}**	**0.02**	**0.13^{a}**	**0.02**
Pretentious	**0.04^{a}**	**0.01**	**0.04^{a}**	**0.01**	**0.08^{a}**	**0.01**	**0.07^{a}**	**0.02**	**0.03^{a}**	**0.01**	**0.06^{a}**	**0.01**	**0.05^{a}**	**0.01**
Relaxed	0.20^{c}	0.02	0.14^{bc}	0.02	0.07^{a}	0.02	0.04^{a}	0.01	0.15^{bc}	0.02	0.10^{ab}	0.02	0.07^{b}	0.01
Sensual	0.06^{a}	0.01	0.04^{a}	0.01	0.02^{a}	0.01	0.02^{a}	0.01	0.06^{a}	0.01	0.04^{a}	0.01	0.04^{a}	0.01
Simple	0.16^{bc}	0.02	0.23^{a}	0.02	0.19^{ab}	0.02	0.12^{bcd}	0.02	0.20^{ab}	0.02	0.12^{bcd}	0.02	0.11^{cd}	0.02
Sophisticated	0.06^{a}	0.01	0.03^{a}	0.01	0.03^{a}	0.01	0.03^{a}	0.01	0.05^{a}	0.01	0.04^{a}	0.01	0.04^{a}	0.01
Traditional	0.17^{b}	0.02	0.15^{b}	0.02	0.05^{a}	0.01	0.04^{a}	0.01	0.09^{a}	0.01	0.04^{a}	0.01	0.04^{a}	0.01
Uninspired	0.08^{c}	0.02	0.13^{bc}	0.02	0.18^{a}	0.02	0.23^{ab}	0.02	0.07^{c}	0.02	0.10^{c}	0.02	0.12^{bc}	0.02

^{abcd}Product mean iStd.IRT values with different letter codes by row (emotion term) are significantly different (*Tukey, $p < 0.05$*).

Table 9
Summary of Chi-square and Pr(>Chi) values from Analysis of Deviance of generalised linear mixed effect models (GLME) on mean citation proportions, and F values and Pr(>F) values from ANOVA of linear mixed effect models (LMM) on IRT mean iStd.IRT of emotion terms.

Emotion term	CATA (GLME) *Product* (df = 6)		IRT (LMM) *Product* (df = 6)	
	Chi-square	Pr > Chi	F value	Pr > F
Adventurous	40.489	**<0.001**	3.021	**0.006**
Bored	34.163	**<0.001**	5.217	**<0.001**
Cheap	3434143.804	**<0.001**	5.319	**<0.001**
Classy	9.384	0.153	2.903	**0.008**
Comforted	118.839	**<0.001**	14.735	**<0.001**
Energised	30.677	**<0.001**	3.225	**0.004**
Feminine	14.954	**0.021**	1.921	0.075*
Genuine	62.638	**<0.001**	6.301	**<0.001**
Happy	93.832	**<0.001**	12.114	**<0.001**
Inspired	30.478	**<0.001**	3.226	**0.004**
Irritated	92.89	**<0.001**	8.637	**<0.001**
Masculine	10.715	0.098*	0.966	0.447
Modern	18.887	**0.004**	1.827	0.092*
Pretentious	45.863	**<0.001**	1.732	0.111
Relaxed	108.215	**<0.001**	12.364	**<0.001**
Sensual	11.514	**0.021**	2.629	**0.016**
Simple	39.135	**<0.001**	7.53	**<0.001**
Sophisticated	14.724	**0.023**	1.319	0.247
Traditional	135.031	**<0.001**	16.17	**<0.001**
Uninspired	96.729	**<0.001**	10.558	**<0.001**

Bold font represents significant *Pr(>Chi) < 0.05* for CATA and *Pr(>F) < 0.05* for IRT. **Pr < 0.10*.

citation proportions are low, those participants may have taken longer to check the emotion. Although this hypothesis needs further investigation, measuring response times in a Go/No-go approach could provide useful data in preliminary studies to help remove redundant terms from lengthy lexicons.

The high negative correlation of response time and citation proportions of CATA also brings into question whether responding under time pressure is more representative of actual consumer decision making and hence the product experience. Some researchers showed association of lower response times with higher accuracy and confidence for visual stimulus (Hagura, Esmaily, & Bahrami, 2023) and recognition/recall memory (Robinson, Johnson, & Herndon, 1997; Weidemann & Kahana, 2016; Zylberberg, Fetsch, & Shadlen, 2016). In the present study, some participants indicated different responses in CATA where there is no time pressure, to in the IRT which is time restricted. It infers they may be switching their responses in the absence of time pressure in CATA, representing a more conscious-explicit choice pathway or hesitancy of their decisions (Walczyk, Roper, Seemann, & Humphrey, 2003) and hence differ from their explicit responses in the Go/No-go task. Additionally, this difference in response was only observed for some emotion terms, independent of product. This also posits whether some emotions require more time to be evoked and given explicit consideration. For example, in multiple sip evaluation of temporal emotion responses, 'nostalgic' was cited more often towards latter sips compared to higher citations given for 'happy' at sip 1 (Weerawarna N. R. P., Godfrey, Ellis, & Hort, 2021b). Also, it is important to investigate whether explicit CATA or IRT represented the actual consumer experience of the product, possibly by exploring future purchase behaviours or associations with other affective and sensory responses of the product set.

Overall, data modelling showed significant *product* effect on more emotion terms in CATA than in IRT, indicating binary format of data in CATA and subsequent higher product discrimination than with continuous data in IRT. Even marginal, it was interesting to note that CATA was differentiating more between cow milks and IRT in the PBMA space. In fact, purportedly IRT is outperforming the traditional CATA for this PBMA product criterion. A few exceptions were that some emotions differentiated by mean comparisons in CATA were not differentiated in IRT and vice versa. The Correspondence Analysis and confidence ellipses indicated that CATA appeared to be more distinct. However, it is important to note that IRT is a continuous measure (Gomez et al., 2007; Pierguidi et al., 2023; Porcherot et al., 2022) versus CATA data which are binary (Ares et al., 2017), and hence raises the question as to whether CATA's binary data format exaggerates product discrimination by ignoring participant indecisiveness revealed in the continuous measure of response time in IRT.

IRT might not be ideal for longer and complex emotion lexicons. According to cognitive psychology, longer lists of terms are reported to reduce consumer attention to the evaluation task resulting satisficing responses (Krosnick, 1999; Krosnick & Alwin, 1987; Rasinski, Mingay, & Bradburn, 1994). Additionally, if the emotion term is complicated to comprehend, time taken to understand the emotion is likely to be recorded as the response time as an indication of association of an emotion to the product. However, an optimum lexicon length or level of complexity is not confirmed for IRT Go/No-go nor emotion CATA, as opposed to a recommended length of up to 20 terms in sensory CATA (Ares, Antúnez, Giménez, & Jaeger, 2015). Therefore, more research is required to explore the optimum lexicon length for IRT Go/No-go and emotion CATA tasks. Further, the complexity of explaining IRT Go/No-go tasks to naïve consumers needs to be weighed against the insights captured from the approach with respect to simplicity and rapidness in CATA (Jaeger & Ares, 2023). Additionally, IRT Go/No-go requires specific software for data collection, as it is not available in current sensory consumer science data collection software.

Optimum approaches for data transformations in reaction time needs further investigations as minimum reaction time can potentially also be considered an outlier, and hence a limitation of this study. However, data analysis of this study was reliant on the central limit theorem eliminating the problems associated with the small risk of an outlier or the distribution of the response variable. If an assessor's minimum response time is an outlier, then that assessor will have less influence on the overall findings than other respondents.

5. Conclusions

In summary, IRT Go/No-go did not outperform CATA on assessor consistency across sessions, and frequency of selecting emotions. Compared to CATA, continuous measure of time in IRT provided additional information on the relevance of selected emotions to a particular product. Product discrimination generally was higher in CATA between cow milks and IRT in PBMA space, but as an artefact CATA appears more distinct based on confidence ellipses. Being under no time pressure to answer in CATA, some participants seemed to change their mind in their selections compared to when they were subjected to time restrictions in IRT Go/No-go and hence questions which approach provides more accurate data on consumer product experience.

It is important to understand which of the two methods, explicit versus implicit, or a combination approach is more representative of actual product experience, and whether some emotions require more time to be evoked and given explicit consideration. Further investigations are needed to validate the findings of this study and provide recommendations on the choice, or combination, of methods for researchers. This can be achieved by evaluating various product matrices with different sensory profiles that can elicit a wider range of valence and arousal of emotions.

Funding

This project was funded by a NZ Ministry of Business and Education (MBIE) Catalyst Grant (70161-CNZSFF-MAU).

CRediT authorship contribution statement

Maheeka Weerawarna N.R.P.: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing – original draft,

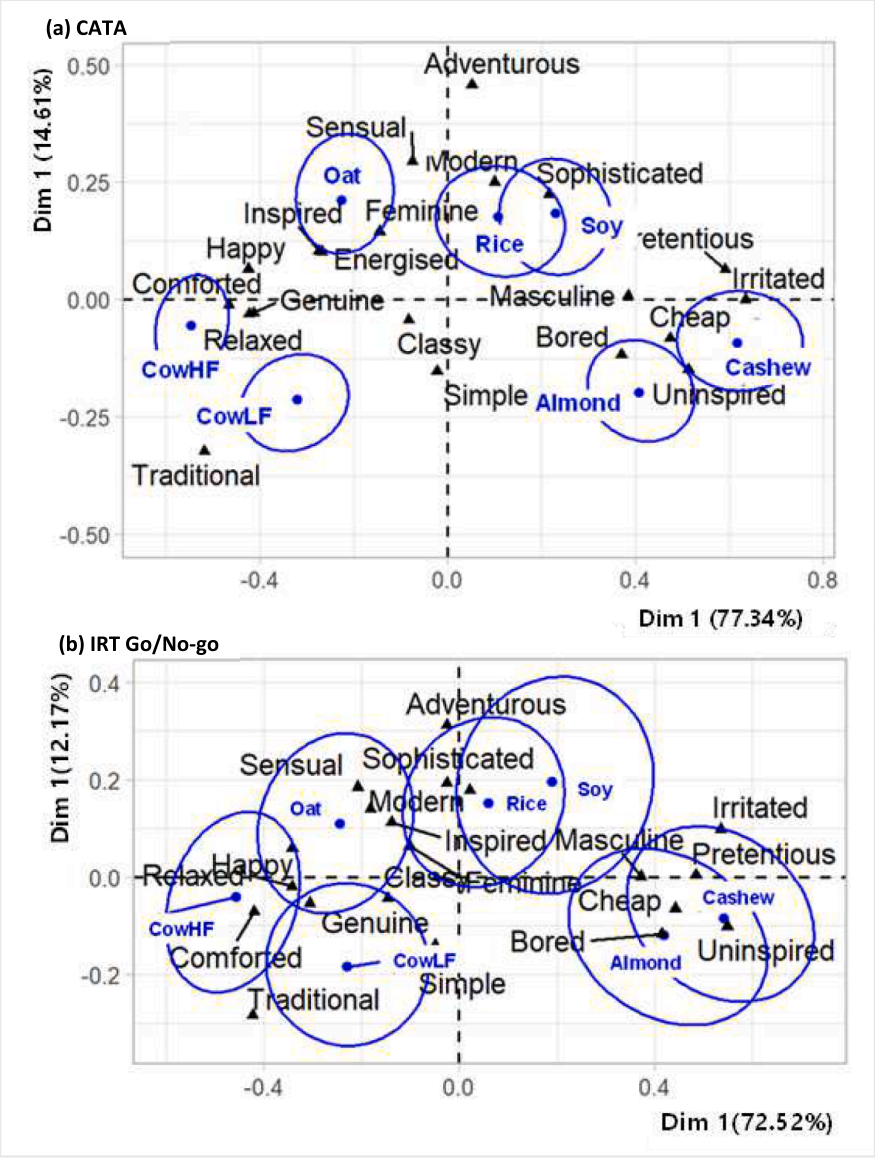


Fig. 6. Biplot of the first two dimensions from Correspondence Analysis with 95 % confidence ellipses (a) mean citation proportions of emotion CATA data, (b) means of iStd.IRT (=1 – Std.IRT) data. Products are plotted as circles and emotion terms as triangles.

Writing – review & editing. **A. Jonathan R. Godfrey:** Methodology, Formal analysis, Writing – review & editing. **Malcolm Loudon:** Methodology. **Meika Foster:** Methodology, Supervision, Writing – review & editing, Funding acquisition. **Joanne Hort:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Funding acquisition, Supervision.

Declaration of Competing Interest

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

Data availability

The data that has been used is confidential.

References

Abdi, H., & Béra, M. (2014). Correspondence analysis. In R. Alhajj, & J. Rokne (Eds.), *Encyclopedia of social networks and mining* (pp. 275–284). New York: Springer Verlag.

Agresti, A. (2018). Generalised linear models. In *An introduction to categorical data analysis* (Third edition ed.): John Wiley & Sons, Inc.

Anzman-Frasca, S., & Ehrenberg, S. (2018). Learning to like: Roles of repeated exposure and other types of learning. In J. C. Lumeng, & J. O. Fisher (Eds.), *Pediatric Food Preferences and Eating Behaviors* (pp. 35–52). Academic Press. https://doi.org/10.1016/B978-0-12-811716-3.00002-6.

Ares, G., Antúnez, L., Giménez, A., & Jaeger, S. R. (2015). List length has little impact on consumers' visual attention to CATA questions. *Food Quality and Preference, 42*, 100–109. https://doi.org/10.1016/j.foodqual.2015.01.015

Ares, G., de Andrade, J. C., Antúnez, L., Alcaire, F., Swaney-Stueve, M., Gordon, S., & Jaeger, S. R. (2017). Hedonic product optimisation: CATA questions as alternatives to JAR scales. *Food Quality and Preference, 55*, 67–78. https://doi.org/10.1016/j.foodqual.2016.08.011

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berridge, K., & Winkielman, P. (2003). What is an unconscious emotion? (The case for unconscious liking). *Cognition and Emotion, 17*(2), 181–211.

Beurteaux, D. (2023). Next-Generation Sensory Science. *Food Technology*.

Bi, J., & Kuesten, C. (2022). Commentary on Meyners and Hasted (2021): On the applicability of ANOVA models for CATA data, Food Quality and Preference, 92.

Food Quality and Preference, 95, Article 104340. https://doi.org/10.1016/j.foodqual.2021.104340
Cardello, A. V., Llobell, F., Giacalone, D., Roigard, C. M., & Jaeger, S. R. (2022). Plant-based alternatives vs. dairy milk: Consumer segments and their sensory, emotional, cognitive and situational use responses to tasted products. *Food Quality and Preference, 104599*. https://doi.org/10.1016/j.foodqual.2022.104599
Cervellon, M. C., Dube, L., & Knauper, B. (2007). Implicit and explicit influences on spontaneous and deliberate food choices. *Paper presented at the Advances in Consumer Research*.
Cliceri, D., Spinelli, S., Dinnella, C., Prescott, J., & Monteleone, E. (2018). The influence of psychological traits, beliefs and taste responsiveness on implicit attitudes toward plant- and animal-based dishes among vegetarians, flexitarians and omnivores. *Food Quality and Preference, 68*, 276–291. https://doi.org/10.1016/j.foodqual.2018.03.020
De Houwer, J., & Moors, A. (2010). Implicit measures: Similarities and differences. *Handbook of implicit social cognition: Measurement, theory, and applications*, 176-193.
de Wijk, R. A., & Noldus, L. P. J. J. (2021). Using implicit rather than explicit measures of emotions. *Food Quality and Preference, 92*, Article 104125. https://doi.org/10.1016/j.foodqual.2020.104125
Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412–431.
Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*(2), 115–141.
Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*(2), 229.
Gavazzi, G., Giovannelli, F., Noferini, C., Cincotta, M., Cavaliere, C., Salvatore, M., … Viggiano, M. P. (2023). Review article: Subregional prefrontal cortex recruitment as a function of inhibitory demand: An fMRI metanalysis. *Neuroscience & Biobehavioral Reviews, 152*, Article 105285.
Greenwald, A. G., & Banji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102* (1) (1995), 4-27.
Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*(3), 389.
Gossard, M. H., & York, R. (2003). Social structural influences on meat consumption. *Human Ecology Review, 10*(1), 1–9.
Hagura, N., Esmaily, J., & Bahrami, B. (2023). Does decision confidence reflect effort? *PLoS One1, 18*(2), e0278617.
Hermans, D., Houwer, J. D., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition & Emotion, 8*(6), 515–533.
Hoogeveen, H. R., Jolij, J., ter Horst, G. J., & Lorist, M. M. (2016). Brain Potentials Highlight Stronger Implicit Food Memory for Taste than Health and Context Associations. *PLoS One1, 11*(5), e0154128.
Horska, E., Bercik, J., Krasnodebski, A., Matysik-Pejas, R., & Bakayova, H. (2016). Innovative approaches to examining consumer preferences when choosing wines. *Agricultural Economics, 62*(3), 124–133.
ISO8589:2007(E). Sensory analysis — General guidance for the design of test rooms.
Jaeger, S. R., & Ares, G. (2023). Using Check-All-That-Apply (CATA) Questions in Emotion Questionnaires. In M. Bensafi (Ed.), *Basic Protocols on Emotions, Senses, and Foods* (pp. 59-72). New York, NY: Springer US. 10.1007/978-1-0716-2934-5_6.
Jaeger, S. R., Chheang, S. L., Yin, J., Bava, C. M., Gimenez, A., Vidal, L., & Ares, G. (2013). Check-all-that-apply (CATA) responses elicited by consumers: Within-assessor reproducibility and stability of sensory product characterizations. *Food Quality and Preference, 30*(1), 56–67. https://doi.org/10.1016/j.foodqual.2013.04.009
Kahneman, D.-F. (2002). *Representativeness Revisited: Attribute Substitution in Intuitive Judgment.* Cambridge, UK: Cambridge University Press.
Kantner, J., & Lindsay, D. S. (2012). Response bias in recognition memory as a cognitive trait. *Memory & Cognition, 40*(8), 1163–1177. https://doi.org/10.3758/s13421-012-0226-0
Kemp, S. E., Nyambayo, I., Rogers, L., Sanderson, T., & Villarino, C. B. (2021). Trends in food sensory science. *Food Science and Technology, 35*(4), 46–50. https://doi.org/10.1002/fsat.3504_13.x
Kostyra, E., Wasiak-Zys, G., Rambuszek, M., & Waszkiewicz-Robak, B. (2016). Determining the sensory characteristics, associated emotions and degree of liking of the visual attributes of smoked ham. A multifaceted study. *LWT - Food Science and Technology, 65*, 246–253. https://doi.org/10.1016/j.lwt.2015.08.008
Kraus, A. A., & Piqueras-Fiszman, B. (2016). Sandwich or sweets? An assessment of two novel implicit association tasks to capture dynamic motivational tendencies and stable evaluations towards foods. *Food Quality and Preference, 49*, 11–19. https://doi.org/10.1016/j.foodqual.2015.11.005
Krieg, E. F., Jr, Chrislip, D. W., Letz, R. E., Otto, D. A., Crespo, C. J., Brightwell, W. S., & Ehrenberg, R. L. (2001). Neurobehavioral test performance in the third National Health and Nutrition Examination Survey. *Neurotoxicology and teratology, 23*(6), 569–589.
Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*(1), 537–567.
Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly, 51*(2), 201–219.
Lagast, S., Gellynck, X., Schouteten, J. J., De Herdt, V., & De Steur, H. (2017). Consumers' emotions elicited by food: A systematic review of explicit and implicit methods. *Trends in Food Science & Technology, 69*, 172–189. https://doi.org/10.1016/j.tifs.2017.09.006
Le, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software, 25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01
Maass, A., Castelli, L., & Arcuri, L. (2000). Social Identity Processes: Trends in Theory and Research. In. London: SAGE Publications Ltd. Retrieved from https://sk.sagepub.com/books/social-identity-processes 10.4135/9781446218617.
Mai, R., & Hoffmann, S. (2017). Indirect ways to foster healthier food consumption patterns: Health-supportive side effects of health-unrelated motives. *Food Quality and Preference, 57*, 54–68. https://doi.org/10.1016/j.foodqual.2016.11.009
Mai, R., Hoffmann, S., Hoppert, K., Schwarz, P., & Rohm, H. (2015). The spirit is willing, but the flesh is weak: The moderating effect of implicit associations on healthy eating behaviors. *Food Quality and Preference, 39*, 62–72. https://doi.org/10.1016/j.foodqual.2014.06.014
Matukin, M., & Ohme, R. (2017). Implicit measures of attitudes in market research. *Ethics and Neuromarketing: Implications for Market Research and Business Practice*, 113-122.
Meyners, M., & Castura, J. C. (2016). Randomization of CATA attributes: Should attribute lists be allocated to assessors or to samples? *Food Quality and Preference, 48*, 210–215. https://doi.org/10.1016/j.foodqual.2015.09.014
Meyners, M., Castura, J. C., & Carr, B. T. (2013). Existing and new approaches for the analysis of CATA data. *Food Quality and Preference, 30*(2), 309–319. https://doi.org/10.1016/j.foodqual.2013.06.010
Ng, M., Chaya, C., & Hort, J. (2013). Beyond liking: Comparing the measurement of emotional response using EsSense Profile and consumer defined check-all-that-apply methodologies. *Food Quality and Preference, 28*(1), 193–205. https://doi.org/10.1016/j.foodqual.2012.08.012
Ng, M., & Hort, J. (2015). Insights into measuring emotional response in sensory and consumer research. In *Rapid Sensory Profiling Techniques* (pp. 71–90). Woodhead Publishing. https://doi.org/10.1533/9781782422587.1.71.
Niedziela, M. M., & Ambroze, K. (2021). The future of consumer neuroscience in food research. *Food Quality and Preference, 92*, Article 104124. https://doi.org/10.1016/j.foodqual.2020.104124
Nijman, M., James, S., Dehrmann, F., Smart, K., Ford, R., & Hort, J. (2019). The effect of consumption context on consumer hedonics, emotional response and beer choice. *Food Quality and Preference, 74*, 59–71. https://doi.org/10.1016/j.foodqual.2019.01.011
Ohme, R., & Boshoff, C. (2019). The role of implicit learning in logo substitution. *Journal of Consumer Marketing, 36*(5), 610–619. https://doi.org/10.1108/JCM-11-2017-2430
Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*(3), 277.
Pellicer-Sánchez, A., & Schmitt, N. (2012). Scoring Yes–No vocabulary tests: Reaction time vs. nonword approaches. *Language Testing, 29*(4), 489–509. https://doi.org/10.1177/0265532212438053
Perry, C. L., McGuire, M. T., Neumark-Sztainer, D., & Story, M. (2001). Characteristics of vegetarian adolescents in a multiethnic urban population1 1The full text of this article is available via JAH Online at http://www.elsevier.com/locate/jahonline. *Journal of Adolescent Health, 29*(6), 406-416. https://doi.org/10.1016/S1054-139X(01)00258-0
Pierguidi, L., Prescott, J., Spinelli, S., Dinnella, C., De Toffoli, A., White, T. L., & Monteleone, E. (2023). Chemosensory quality and intensity reflected in implicit affective responses. *Food Quality and Preference, 104857*. https://doi.org/10.1016/j.foodqual.2023.104857
Porcherot, C., Raviot-Derrien, S., Beague, M.-P., Henneberg, S., Niedziela, M., Ambroze, K., & McEwan, J. A. (2022). Effect of context on fine fragrance-elicited emotions: Comparison of three experimental methodologies. *Food Quality and Preference, 95*, Article 104342. https://doi.org/10.1016/j.foodqual.2021.104342
R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.
Rasinski, K. A., Mingay, D., & Bradburn, N. M. (1994). Do respondents really "mark all that apply" on self-administered questions? *Public opinion quarterly, 58*(3), 400–408.
Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *114*, 510-532. 10.1037/0033-2909.114.3.510.
Robinson-O'Brien, R., Perry, C. L., Wall, M. M., Story, M., & Neumark-Sztainer, D. (2009). Adolescent and Young Adult Vegetarianism: Better Dietary Intake and Weight Outcomes but Increased Risk of Disordered Eating Behaviors. *Journal of the American Dietetic Association, 109*(4), 648–655. https://doi.org/10.1016/j.jada.2008.12.014
Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology, 82*(3), 416.
Rothgerber, H. (2014). A comparison of attitudes toward meat and animals among strict and semi-vegetarians. *Appetite, 72*, 98–105. https://doi.org/10.1016/j.appet.2013.10.002
Shelton, J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, memory, and cognition, 18*(6), 1191.
Soulikhan, F., Norré, B., & Cheikhrouhou, N. (2021). *Implicit Measurement Method for Consumer Engagement in Online Brand Communities*.
Spinelli, S., & Monteleone, E. (2018). Emotional responses to products. In G. Ares, & P. Varela (Eds.), *Methods in Consumer Research* (Volume 1, pp. 261–296). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-102089-0.00011-X.
Tempere, S., Pérès, S., Espinoza, A. F., Darriet, P., Giraud-Héraud, E., & Pons, A. (2019). Consumer preferences for different red wine styles and repeated exposure effects. *Food Quality and Preference, 73*, 110–116. https://doi.org/10.1016/j.foodqual.2018.12.009
Thompson, P., Colebatch, J., Brown, P., Rothwell, J., Day, B., Obeso, J., & Marsden, C. (1992). Voluntary stimulus-sensitive jerks and jumps mimicking myoclonus or pathological startle syndromes. *Movement disorders: official journal of the Movement Disorder Society, 7*(3), 257–262.

Verhulst, F., Hermans, D., Baeyens, F., Spruyt, A., & Eelen, P. (2006). Determinants and predictive validity of direct and indirect measures of recently acquired food attitudes. *Appetite, 46*(2), 137–143. https://doi.org/10.1016/j.appet.2005.11.004

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology, 17*(7), 755–774. https://doi.org/10.1002/acp.914

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of breif measures of positive and negative affects: The PANAS scales. *Journal of Personality and Socila Pshychology, 54*, 1063–1070.

Weerawarna, N. R. P., Godfrey, A. J. R., Ellis, A., & Hort, J. (2021a). Comparing temporal sensory product profile data obtained from expert and consumer panels and evaluating the value of a multiple sip TCATA approach. *Food Quality and Preference, 89*, Article 104141. https://doi.org/10.1016/j.foodqual.2020.104141

Weerawarna N. R. P., M., Godfrey, A. J. R., Ellis, A., & Hort, J. (2021b). *Evaluating temporal multiple sip approaches to characterise product experience.* (Doctor of Philosophy), Massey University.

Weidemann, C. T., & Kahana, M. J. (2016). Assessing recognition memory using confidence ratings and response times. *Royal Society Open Science, 3*(4), Article 150670. https://doi.org/10.1098/rsos.150670

Werle, C. O. C., Trendel, O., & Ardito, G. (2013). Unhealthy food is not tastier for everybody: The "healthy=tasty" French intuition. *Food Quality and Preference, 28*(1), 116–121. https://doi.org/10.1016/j.foodqual.2012.07.007

Wickham, H., François, R., Henry, L., & Müller, K. (2020). dplyr: A Grammar of Data Manipulation (Version R package version 1.0.0). Retrieved from https://CRAN.R-project.org/package=dplyr.

Winkielman, P., Berridge, K. C., & Wilberger, J. (2000) Unconcsious affect for doing without feeling: Subliminal facial expression altrer human consumption. *Unpublished manuscript.*

Winkielman, P., Zajonc, R. B., & Schwarz, N. (1997). Subliminal affective priming resists attributional interventions. *Congnition & Emotion, 11*, 433–465.

Wixted, J. T. (2020). The forgotten history of signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(2), 201–233. https://doi.org/10.1037/xlm0000732

Woodward, H. E., & Treat, T. A. (2015). Unhealthy how?: Implicit and explicit affective evaluations of different types of unhealthy foods. *Eating Behaviors, 17*, 27–32. https://doi.org/10.1016/j.eatbeh.2014.12.011

Worsley, A., & Skrzypiec, G. (1998). Teenage Vegetarianism: Prevalence, Social and Cognitive Contexts. *Appetite, 30*(2), 151–170. https://doi.org/10.1006/appe.1997.0118

Zylberberg, A., Fetsch, C. R., & Shadlen, M. N. (2016). The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *eLife, 5*, e17688.