



Contents lists available at ScienceDirect

# Food Quality and Preference

journal homepage: [www.elsevier.com/locate/foodqual](http://www.elsevier.com/locate/foodqual)

## Effects of intrinsic and extrinsic product characteristics related to protein source, health and environmental sustainability, on product choice and sensory evaluation of meatballs and plant-based alternatives

Caroline Giezenaar<sup>a</sup>, A. Jonathan R. Godfrey<sup>b</sup>, Meika Foster<sup>c</sup>, Joanne Hort<sup>a,d,\*</sup>

<sup>a</sup> 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

<sup>b</sup> Statistics Group, School of Mathematical and Computational Sciences, Massey University, Palmerston North, New Zealand

<sup>c</sup> Edible Research Limited, Ohoka 7475, New Zealand

<sup>d</sup> Riddet Institute, Massey University, Palmerston North 4410, New Zealand

### ARTICLE INFO

#### Keywords:

Product choice  
Consumer perception  
Plant-based meat alternatives  
Environmental sustainability  
Health  
Intrinsic and extrinsic product attributes

### ABSTRACT

Health is cited as a major driver for substituting meat with plant-based meat alternatives (PBMA). In contrast, many consumers are reportedly unaware of the difference in environmental impact between meat and PBMA. This study determined relative effects of internal (protein source) and external (health/origin/sustainability) product attributes on product choice and (expected) liking of meat and PBMA.

Participants ( $n = 144$ ) ranked ten sets of three product concepts from best to worst in a discrete choice experiment (DCE). Random combinations of four product attributes were presented on hypothetical product packages: i) protein source (beef/chickpea/soy protein); ii) health-star rating (1.5/4.5 stars); iii) ingredient origin (local/imported); iv) sustainability rating (green tick absent/present), and product worth, related to each of the 24 possible attribute combinations, was determined. Secondly, commercially-available beef, soy, and chickpea products were sensorially evaluated for expected/actual liking, both blindly and alongside combinations of internal/external product attributes used in the DCE.

The DCE indicated all product attributes affected product worth. However, two consumer segments were established: i) a 'Protein Cluster' (47%) comprising consumers who chose 'beef' products regardless of the remaining attributes present, and ii) a 'Health/Sustainability Cluster' (53%) of consumers whose choices were driven by positive health/sustainability claims. Protein source drove both *expected* and *actual* liking of the samples ( $p < 0.001$ ). Health/origin/sustainability also contributed to *expected* liking ( $p < 0.05$ ), but only health consistently affected *actual* liking ratings ( $p < 0.05$ ).

In conclusion, half of consumers chose products based on health/origin/sustainability claims, whereas the other half selected products based on protein source. Origin/sustainability claims affected product choice, but minimally affected product experience of plant-based samples.

### 1. Introduction

Plant-based meat alternative (PBMA) sales are increasing rapidly; for example, dollar sales in the US grew 74% between 2019 and 2021 (Good Food Institute, 2022). PBMA are generally classifiable into two categories: those developed to mimic a meat product and those not necessarily produced to imitate the flavour and texture of meat (Macdiarmid

2022). Products that imitate meat (sometimes known as meat analogues) are often made from processed plant-based proteins, such as soy or pea protein isolate, used to achieve a more meat-like texture and a protein content similar to meat (Tso et al. 2021). In contrast, products that do not imitate a meat product are often made from wholefoods, such as legumes, vegetables, or grains. Consumer studies have shown that some consumers prefer meat-like PBMA, whereas other consumers

**Abbreviations:** AIC, Akaike Information Criterion; A-NZ, Aotearoa New Zealand; CE, coefficient estimate; DCE, discrete choice experiment; MS, mean square; PBMA, plant-based meat alternative.

\* Corresponding author.

E-mail address: [j.hort@massey.ac.nz](mailto:j.hort@massey.ac.nz) (J. Hort).

<https://doi.org/10.1016/j.foodqual.2023.105070>

Received 4 August 2023; Received in revised form 27 November 2023; Accepted 4 December 2023

Available online 8 December 2023

0950-3293/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

do not (Collier et al. 2021; Elzerman Johanna et al. 2013; Kerslake et al. 2022; Michel et al., 2021a). These preferences may be affected by the diet of the consumer. For example, in focus groups with New Zealand consumers, omnivores expressed that they were impressed with PBMA that closely resembled meat, whereas vegans and vegetarians felt meat substitutes that closely resembled meat were ‘too real’ or ‘strange’ such that it created negative experiences (Kerslake et al. 2022). Furthermore, meat avoiders and flexitarians have more positive perceptions of PBMA than omnivores in terms of taste, texture, effort to prepare, protein content, and environmental friendliness (Michel et al., 2021b).

Despite increasing availability, the uptake of PBMA overall is still low: an approximate 1.4% market share of the US meat industry was reported in 2021 (Good Food Institute, 2022). Although many PBMA are developed to imitate a meat product, such as burgers, sausages, chicken pieces, minced meats or meatballs, consumers who do not exclude meat from their diet tend to prefer the texture and flavour of ‘real’ meat over PBMA (Slade 2018; Van Loo et al. 2020).

Furthermore, although life cycle analyses have reported that production of the Beyond Burger produces 90% less greenhouse gases and requires 46% less energy, 93% less land, and 9% less water compared to a conventional beef patty (Heller and Keoleian 2018), many consumers are not convinced that PBMA are actually better for health or environmental sustainability. For example, in a recent study 534 Swiss general consumers were asked to rate healthiness and environmental impact of meat and PBMA (Hartmann et al. 2022). On average, meat products were rated to be healthier than PBMA, despite their similar nutrient profiles. Likewise, respondents estimated the environmental impact of a PBMA to be greater, and that of meat to be smaller, than the actual impact: the Beyond Burger was falsely perceived to have a worse environmental impact than a beef entrecôte. Negative perceptions of the healthiness and environmental impact of PBMA may be related to the processed nature of these products: there is consumer perception that minimally processed, or ‘natural’, foods are more environmentally sustainable than processed, or ‘unnatural’, foods (Román et al. 2017). Such perceptions, although not necessarily accurate, may deter consumers from selecting PBMA over meat products.

For PBMA consumers, perceived healthiness is a major driver of product choice, alongside intrinsic product characteristics such as taste, convenience, and appearance (Onwezen et al. 2021; Siegrist and Hartmann 2019; Tso et al. 2021). Environmental sustainability also is often discussed as a driver for PBMA selection, but environmental motivators appear to be generally less important than health motivators in influencing PBMA product choice – at least when choosing hybrid meat/vegetable products (Lang, 2020). Furthermore, ingredient origin may have a large effect on consumer perceptions of environmental impact. For example, imported meat products were perceived as very environmentally unfriendly, whereas local meat products were not, and found that country of origin was the most important indicator of environmental friendliness (Lazzarini et al., 2016). Messages on health and environmental benefits of pulses increased expected and experienced liking of chickpea and black bean spreads (Henn et al., 2023). How the presence of claims regarding health, origin, and environmental benefits on a product package affect perceptions of PBMA, relative to conventional meat, is unknown.

This study aimed to quantify how the presence of product attributes relating to the main product ingredient, and claims regarding health, origin and environmental sustainability affected product choice and acceptance of PBMA, and a meat product as a comparator. It also sought to identify any consumer segmentation based on product attributes driving their product choice. The study used digital immersive environments to optimise the ecological validity of the experiment (Giezenaar and Hort 2021; Liu et al. 2017), where an immersive context of a supermarket was used to represent a point-of-purchase scenario in which the participants completed a product choice experiment, and a kitchen context was used to represent a consumption scenario in which participants evaluated the samples for acceptance.

## 2. Material and methods

This study was evaluated via the Massey Human Ethics Committee process and was judged as low risk (Application ID 4000025647). The experiment followed the storyline of a consumer selecting a meat/veggie-ball product in a supermarket context (Phase 1), then evaluating the products for liking in a home context (Phase 2).

### 2.1. Participants

Participants (n = 144) were recruited using the Food Experience and Sensory Testing (Feast) consumer database, internal emails, and flyer distribution on the Manawatū Campus of Massey University, New Zealand. Inclusion criteria were: aged 18–55 years, willing to consume meat and PBMA, not pregnant or lactating, and not allergic to sample ingredients, and were assessed in a screening survey. Gender, age, self-reported diet (omnivore or flexitarian), and meat attachment scores (Graça et al. 2015) were also collected to determine whether these demographics affected study outcomes. Each participant took part in one session (~45 min) attended by up to eight participants. Participants gave written informed consent and were offered a supermarket voucher to compensate for their time.

### 2.2. Protocol

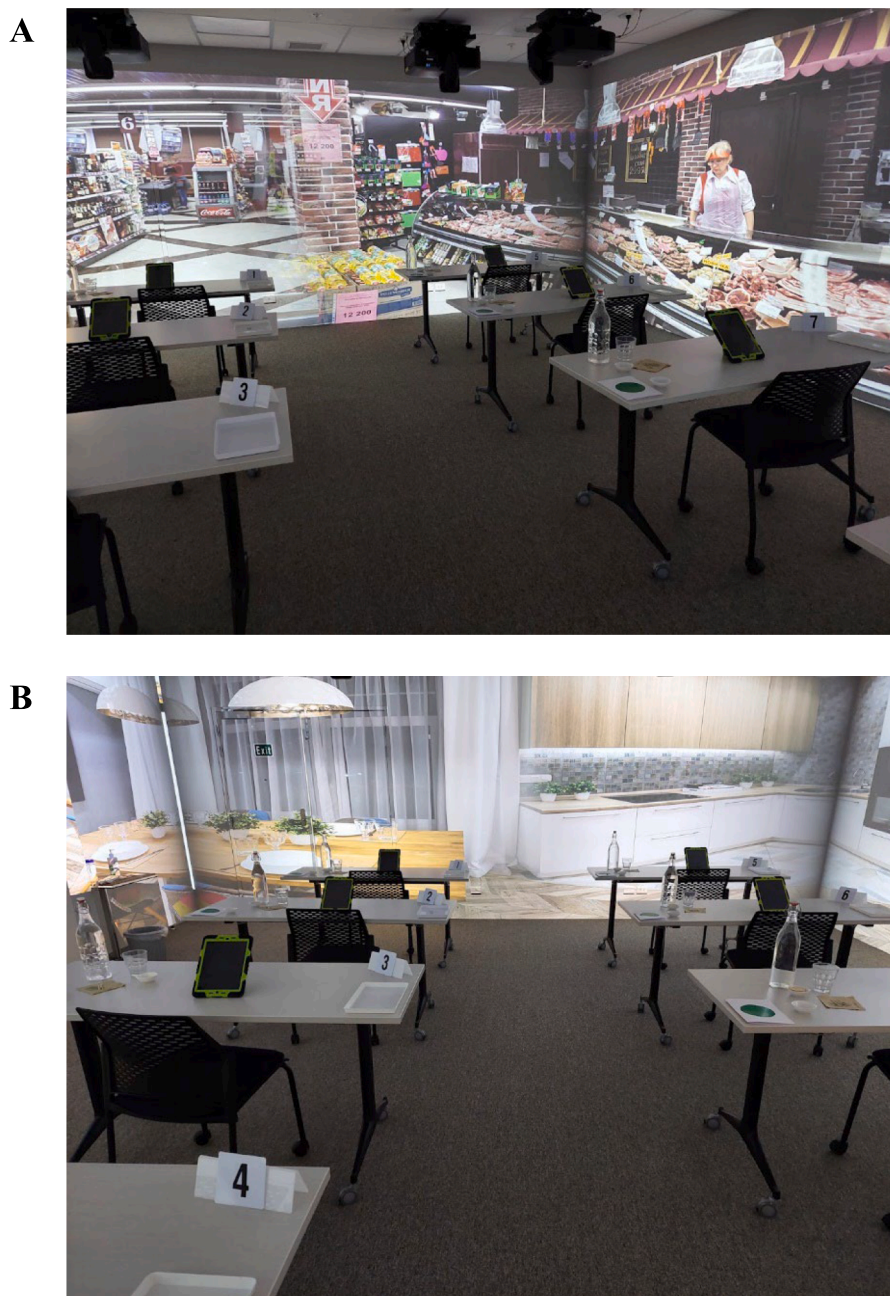
The study was conducted in the ‘Immersive Space’ at the Feast laboratory, where selected 360-degree photographic footage was projected on all four walls through eight projectors (Igloo Vision Ltd., Thornbury, VIC, Australia) and associated audio played through a speaker system. Each participant was seated at a separate table (Fig 1).

#### 2.2.1. Study product attributes and attribute levels

Meat/veggie-balls were selected as the test product as they are a common product in Aotearoa New Zealand (A-NZ) food culture and can be eaten on their own, allowing sample tasting to be relevant without needing to add other ingredients. To determine relative importance of key drivers of product choice and acceptance, four product attributes, with two or three levels each, were selected (Table 1) and manipulated throughout the experiment. Firstly, protein source was selected as an internal attribute, with three levels representing 1) processed plant-based ingredients often used for meat analogues (soy protein isolate), 2) plant-based ingredients used in products that do not imitate meat, often made from legumes (chickpeas), and 3) a meat level (beef) selected as a reference product. The protein attribute levels were selected based on ingredients that are commonly used for plant-based alternatives for meatballs - in terms of shape - that both did and did not mimic meat.

Secondly, since health/nutrition is an important driver of product choice for PBMA consumers, low and high health star ratings (1.5 and 4.5 stars, respectively) were included to reflect whether a product concept was healthy (extrinsic attribute) or not.

Since consumers are reportedly often unaware of the environmental impact of meat and PBMA, two extrinsic attributes relating to the environment were selected to determine whether their presence affected product choice. A statement indicating whether the product was made from NZ (i.e. local) or imported ingredients was added to reflect product origin. A sustainability tick (Green Tick) was used to indicate whether the product contributed to environmental sustainability or not (tick present or absent). Whereas protein source, health, and origin of ingredients included two or three visible attribute levels, the sustainability attribute was only visible when the sustainability tick was present (i.e. its absence was not indicated on the product concepts). Attribute levels were presented via hypothetical packaging of a meat/veggie-ball product created by the researchers presented on an iPad (Apple Inc., California, United States of America) screen. On packaging, product names were ‘Meatballs’ for the beef product, and ‘Veggieballs’ for the plant-based products.



**Fig. 1.** A: The Discrete Choice Experiment took place in a digitally recreated supermarket context to represent a point of purchase. B: The sample evaluation took place in a digitally recreated home context, to represent an eating environment.

### 2.2.2. Determining effects of product attributes on product choice at point of purchase

In Phase 1, participants completed a discrete choice experiment (DCE), which was adopted as an appropriate approach to investigate preference for combinations of product attributes related to nutrition/health and environmental sustainability of meat products and plant-based meat alternatives.

The DCE phase was conducted in a digital supermarket context to represent the product-choice aspect of the study. A supermarket image was projected (Fig. 1A), and supermarket sounds (beeps from product scanning, chatter) (VIABEL, 2019) was audible. The image included aisles of a supermarket, with a meat display, refrigerator aisle, and a dry foods aisle. Participants were verbally instructed that they were grocery shopping for dinner with their family/flat mates/friends, or whoever they would have dinner with usually. They were told they were

shopping for meatballs, or a plant-based alternative for meatballs, and to select their most and least liked product for that occasion between the three products presented to them in each set.

The DCE was hosted through Sawtooth software (Lighthouse Study, Version 9.14.1, Sawtooth Software Inc., Sunny Valley, ID, United States). Using iPads, participants were presented with a set of three product concepts on-screen, each concept with different combinations of each of the four product attributes levels (3 protein attribute levels\*2 health attribute levels\*2 origin attribute levels\*2 sustainability attribute levels = 24 product attribute combinations; Table 1). To determine attribute combinations for each product concept, Sawtooth uses randomised subsets of the full-choice design, while ensuring level balance and near-orthogonality within each respondent's survey. This approach avoids systematic correlations among interactions inherent in fixed designs so that both main effects and higher-order interactions can be

**Table 1**

Product intrinsic (protein), extrinsic (health, origin of ingredients, environmental sustainability) attributes and associated levels.

Product attribute	Attribute levels
Protein source	1. Beef 2. Soy protein isolate 3. Chickpeas
Health	1. 1.5 health star rating 2. 4.5 health star rating 3. Attribute absent <sup>a</sup>
Origin of ingredients	Made from imported ingredients Made from NZ ingredients Attribute absent <sup>a</sup>
Environmental sustainability	Sustainability tick absent Sustainability tick present Attribute absent <sup>a</sup>

<sup>a</sup> Attribute absent level was only applied in Phase 2 of the study (sample evaluation).

determined. Balanced overlap between levels was applied. Participants selected their 'best' and 'worst' choice option between the three product concepts presented. This task was repeated 10 times, each time with different combinations of product concepts. To assess intra-participant reliability, two additional fixed tasks were completed, in which two identical sets of three product concepts were presented within and between all participants (Table 2). The fixed tasks were selected so that there was overlap between the product concepts, reflecting real in-market decisions. Included concepts were selected by the researchers so that they were not equally attractive, to avoid equal shares of preference that could be predicted by a random simulator. Instead, shares of preferences were judged to be ~ 50/30/20%, where one option was expected to be most favoured (Concept 1) and one option was expected to be least favoured (Concept 3) (Omre 2014).

### 2.2.3. Determining effects of product attributes on expected and actual liking ratings

After the DCE, participants left the Immersive Space, and took a seat in reception for 5 min. Meanwhile, the immersive space was set up with the digital context for Phase 2, giving participants the opportunity to 'reset' between the two. In Phase 2, participants evaluated samples of meatballs and plant-based alternative for meatballs, alongside the presentation of the product attributes that were used in Phase 1, on expected and actual liking ratings.

During Phase 2, an image of a home, with the participants facing the kitchen and dining table, was projected (Fig. 1B), and kitchen sounds were played (Cozy Corner, 2020), to represent a consumption scenario. Participants were verbally instructed that they were having dinner at home with their family/flat mates/friends or whoever they would have dinner with usually. They were told that the samples presented were a component of their meal.

**2.2.3.1. Sample selection and presentation.** All samples for Phase 2 were commercially available in A-NZ and included beef meatballs (Gourmet meatballs, Angel Bay, ANZCO Foods Ltd, Christchurch, NZ), a pea/soy protein isolate meatball-analogue (Succulent Plant-Based Meatballs, Birds Eye, Simplot Australia Pty Ltd, Mentone, Australia) and a falafel

**Table 2**

Fixed task included twice in the discrete choice experiment to assess intra-participant reliability.

Fixed task	Concept 1	Concept 2	Concept 3
Beef		Soy protein isolate	Soy protein isolate
4.5 health stars		4.5 health stars	1.5 health stars
Imported ingredients		NZ ingredients	Imported ingredients
Sustainability tick absent		Sustainability tick absent	Sustainability tick present

made from chickpeas (Danny's Falafel, Pita Bread Ltd, Auckland, NZ) which did not imitate a meat product. To select the samples, two meatball products, two plant-based meatball mimics, and two falafel products were bench tested by five staff members of the Feast laboratory. The samples for the study were selected based on palatability, and based on flavour neutrality (i.e., one of the plant-based meatball mimics was deemed to have a strong flavour fennel flavour therefore a more neutral tasting sample was selected).

Samples were cooked immediately prior to the session according to manufacturer recommendations and kept warm (50 °C) in a food warmer (E84 Food Warmer, Barbar, Moffat Ltd, Rolleston, NZ) until serving. Samples were presented monadically, wrapped in aluminium foil labelled with a random three-digit code, on a white ceramic dish. Although evaluated multiple times throughout the experiment, samples were coded differently for each presentation. Participants received half a meat/veggie-ball per serving, equivalent to six meat/veggie-balls over the whole experiment. A mandatory minimum 1 min break was enforced between each sample to allow the participants to cleanse their palate with filtered water (15 °C) and plain gluten-free wafer (OB Finest Gluten Free Wafers Original, Gourmet Food Operations Pty Ltd, Dandenong South, VIC, Australia).

**2.2.3.2. Design of samples sets during sample evaluation phase.** To optimally determine main and interaction effects of the product attributes on expected and actual liking ratings, each product attribute needed to be tested on its own, as well as in combination with other product attributes. Hence, in contrast to Phase 1 in which all four product attributes were present on each product concepts presented to participants, in Phase 2 product concepts with between one to four product attributes were evaluated. This resulted in 81 potential product attribute combinations to account for each combination where a protein source was always present, and remaining attributes were either absent or present in either of their two attribute levels (3 protein attribute levels\*3 health attribute levels\*3 origin attribute levels\*3 sustainability attribute levels = 81 product attribute combinations; Table 1). As 81 attribute combinations for each participant was deemed too many to evaluate within practical or financial constraints, a fractional factorial design was selected for presentation orders of samples (protein source) and product attributes, where each participant was presented with a subset of nine of the 81 combinations. During sample evaluation, participants were presented with five sample sets as depicted in Table 3. Participants evaluated each sample multiple times throughout the session, first blinded (Set 1), then with an increasing number of product attributes presented via onscreen product concepts with a layout identical to those used in the DCE (Sets 2–5). Participants were not informed they were tasting the same sample multiple times. When a packaging concept was presented (Sets 2–5), samples were evaluated for *expected* liking based on the product package alone, and *actual* liking based on both the product package and sample sensory experience, to determine whether expectations based on a product package met the actual product experience. Data for sample evaluation were collected using Compusense® Cloud Software (CompuSense Inc., Ontario, Canada) on iPads (Apple Inc., California, United States of America). Liking was measured on a labelled affective magnitude (LAM) scale (Schutz and Cardello 2001) ranging from greatest imaginable dislike to greatest imaginable like.

In Set 1 the beef, soy and chickpea samples were presented monadically in a blind tasting. Each sample was rated for liking. For Set 2, the beef, soy, and chickpea samples were presented monadically and with on-screen product packaging informing them of the protein source. All three samples were then assessed again by a participant across Sets 3–5 but only one protein source was sampled in duplicate per Set. In Set 3, one of the remaining attributes, (i.e. health star rating, origin of ingredients, or sustainability tick), was introduced via addition to the on-screen packaging, with no information regarding the other two attributes present. In Set 4, a second attribute was added, and finally in Set 5,

**Table 3**

Schematic overview of the sample evaluation phase for each individual participant with packaging that matched the protein source.

	Set 1	Set 2	Set 3	Set 4	Set 5
Product evaluated <sup>a</sup>	A B C	A B C	A A	B B	C C
On-the-pack product attributes presented:	None (blind evaluation)	1. Protein source only (beef, soy protein isolate, chickpea)	1. Protein source, 2. 1 of 3 external attributes (Health, origin OR sustainability)	Protein source, the attribute presented in Set 3 one other attribute	Protein source All attributes
Attribute levels presented	N/A	N/A	The two samples within each set displayed opposite attribute levels		
Sample presentation design	Balanced design within Set 1	Balanced design within Set 2	Fractional, balanced design between Sets 3–5		

<sup>a</sup> Letters A, B, and C reflect the nature of samples, i.e. a combination of different letters represents that different protein source samples were presented (Set1-2) whereas a combination of the same letter within a Set represents that a protein source sample is presented in duplicate (Set 3–5). Since the design of the study is fractional, A, B, and C can reflect any sample (i.e. beef, soy, or chickpea).

all three attributes were present. Within each Set, samples were presented according to a balanced design. As the protein attribute was tied to the sample presented, the protein attribute remained present throughout Sets 2–5. Within Sets 3–5, each sample was evaluated in duplicate but the packaging presented contained contrasting product attribute levels to optimally determine interaction effects. For example, if the first product in Set 4 had 4.5 health stars and was made from imported ingredients, the second product in Set 4 automatically had 1.5 health stars and was made from NZ ingredients (example in Table 4). The order of protein source presentation across Sets 3–5, along with which product attributes were presented, was determined with the use of a Williams Latin square in which each combination of protein source presentation order and extrinsic attribute (health stars, origin, sustainability) presentation order appeared once (36 possibilities). Sample order presentation within each Set was determined using separate Williams Latin Squares for each Set, and was balanced between participants.

### 2.3. Statistical analyses

Statistical analyses were performed using R version 4.2.0 through R Studio software version 2022.2.3.492 (R Core Team 2022) with  $\alpha = 0.05$ . Package dplyr (Wickham et al. 2022) was used for data handling.

#### 2.3.1. Participant sample size based on data collection plan allowing for main and interaction effects

A data collection plan and scripts of models to run data analyses was created in R notebook (R Core Team 2022) and applied using fictitious response data to ensure evaluation of up to four-way interactions of product attributes on (expected) liking ratings was possible. Based on analyses of this fictitious data with increasing participant numbers, it was determined that at least 120 participants were required to run the statistical four-way mixed effects models, using the correct number of degrees of freedom. Consequently 144 participants were recruited to ensure sufficient data points, and so that each of the 36 protein-attribute combinations were repeated four times.

#### 2.3.2. Determining intra-participant consistency in the discrete choice experiment

To determine intra-participant consistency, the kappa statistic (Meyer et al. 2023) was calculated based on analyses of a cross-tabulation of the rankings of the three products presented within each fixed task. The higher the intra-participant consistency between the fixed tasks, the higher the validity and accuracy of the data collection. Responses to the fixed tasks were not included in the remaining analyses.

#### 2.3.3. Determining drivers of product choice in the discrete choice experiment

Sawtooth provides absolute utility scores of each product attribute

level for each individual participant computed with Hierarchical Bayesian modelling using the Sawtooth choice-based conjoint hierarchical Bayes (CBC/HB) System (Sawtooth Software Inc. 2021). By design, the sum of the utility scores for all levels within an attribute is zero. To create a measure of sensitivity independent of the number of levels for each attribute, ranges of utility scores were calculated for each participant by subtracting the minimum utility score from the maximum utility for each attribute. A low range indicates relative ambivalence towards an attribute, whereas a high range indicates that the attribute has a relatively higher importance for a particular participant. Cluster analysis (k-means) was performed on the ranges of utility scores, to determine whether participants could be segmented based on their preferred product attributes. The optimal number of clusters was determined using the elbow and silhouette methods (Kassambara and Mundt 2020). Differences in proportions of gender and diet between each cluster and the total sample were investigated with Chi-square goodness of fit tests. Utility scores range differences between clusters were determined by independent t-tests.

'Best' and 'worst' product concepts in each DCE product set were converted to rankings, with the non-selected product given the middle rank position. Due to the fractional design, the PlackettLuce package (Turner et al. 2020), which enables analysis of fractional designs, was used to convert rankings of product concepts within each set of three from the DCE into importance of each attribute. Firstly, data were organised into 24 columns (1 for each attribute combination included in the DCE experiment) and a single row for each set of three product concepts evaluated by the participants (i.e. 10 sets\*144 participants). Each concept within a set received a score of 1 (most liked product concept), 2 (neither most or least liked product concept), or 3 (least liked product concept) with the remaining product concepts receiving a score of 0 to indicate they were not evaluated. The 'as.rankings' function was applied to convert the matrix product choices into a rankings object. Plackett-Luce modelling was conducted to determine the log coefficient estimates (CE) for each individual product concept, as a measure of product worth (relative preference) for each of the 24 product concepts, where a higher coefficient represents a higher product worth. The lowest ranked product was given a coefficient of 0, and the coefficient estimates of the remaining products were relative to the lowest rated product. A linear model with the product worth as response variable, and the attributes as fixed factors, was then used to determine effects of protein source, health star rating, origin of ingredients, and presence of the sustainability tick on the coefficient estimates. The optimal model to predict product coefficients was determined using a forward stepwise algorithm from the baseline model. The baseline model included main effects of the four attribute factors, while the stepwise algorithm identified important interactions between the attributes. The Akaike Information Criterion (AIC) was used to compare best fit of the data between models.

The variation each attribute caused in the model was assessed by

**Table 4**

An example of the on-the-pack product attributes presented to a participant in Sample Sets 1–5.

	Set 1	Set 2	Set 3	Set 4	Set 5
Sample 1 No information presented					
Sample 2 No information presented					
Sample 3 No information presented			N/A – Set 3 contained only two samples	N/A – Set 4 contained only two samples	N/A – Set 5 contained only two samples

determining the mean square (MS) of each attribute factor within the model, where higher variation indicates relatively higher importance of the respective attribute. Protein source was the only attribute with three levels. To determine whether a particular protein source (i.e. beef, soy or chickpea) was responsible for a large part of the variation (i.e. one protein was much more negatively or positively perceived than the other two), a sensitivity analysis was undertaken whereby product concepts of one protein source were excluded from the analyses one by one. Mean squared values for each possible protein pair were compared to the mean square of protein factor in the overall model (including all protein sources).

**2.3.4. Determining effect of presentation order, protein source, and product attributes on sample expected and actual liking ratings**

To determine effects of presentation order and protein source on sample liking ratings in Set 1 and 2, mixed effect models with participants as a random effect, and presentation order number protein source as a fixed effects. To compare actual liking ratings in Set 2 with liking ratings in Set 1 and expected liking ratings in Set 2 within each protein source, paired t-tests were performed. To determine the effects of product attributes and participant characteristics on expected and actual liking ratings, mixed effects models were performed on the expected and actual sample ratings separately for each protein source, with participants as a random effect, and health star rating, origin of ingredients and presence of a sustainability tick as fixed factors, along with gender, age, diet, meat attachment score, and cluster membership. Interactions

between health, origin and sustainability were allowed, however, higher-level interactions were removed from the model if not supported by their underlying lower-level interactions. If a main effect was significant, differences between attribute levels were determined with estimated marginal means post hoc testing, with Benjamini Hochberg corrections for multiple comparisons.

To determine how much ‘actual’ liking related to ‘expected’ liking ratings, r-squared values, as a measure of model predictability, of four mixed effects models predicting actual liking ratings were determined. These models all included participants as a random effect, and included i) a null model without any fixed factors; ii) a model with the protein, health, origin, and sustainability product attributes as fixed factors; iii) a model with ratings of expected liking as a fixed factor; and iv) a model with both product attributes and expected liking ratings as fixed factors. The four models were run to determine whether actual liking was better predicted by expected liking or the mere presence of the attributes, so that the relative importance of perception could be distinguished from the reality of the treatment levels.

### 3. Results

#### 3.1. Participants

The study was completed by 144 participants (39 men and 105 women) with a median age of  $37 \pm 1$  y. Omnivores ( $n = 73$ ; 20 men) had higher meat intakes ( $8.8 \pm 2.8$  meals per week;  $p = 0.001$ ) compared to flexitarians ( $n = 71$ ; 19 men; meat, mean  $\pm$  SD:  $7.3 \pm 2.7$  meals per week). Omnivores scored consistently higher on all components of the meat attachment questionnaire: hedonism (Mean  $\pm$  SD: omnivores:  $4.0 \pm 0.7$ ; flexitarians:  $3.5 \pm 0.07$ ;  $p < 0.001$ ), affinity (omnivores:  $3.0 \pm 0.8$ ; flexitarians:  $2.7 \pm 0.8$ ;  $p = 0.03$ ), entitlement (omnivores:  $3.3 \pm 0.7$ ; flexitarians:  $2.9 \pm 0.8$ ;  $p = 0.003$ ), and dependence (omnivores:  $3.1 \pm 0.5$ ; flexitarians:  $2.7 \pm 0.5$ ;  $p < 0.001$ ), and total mean scores (omnivores:  $3.2 \pm 0.4$ ; flexitarians:  $2.9 \pm 0.4$ ;  $p < 0.001$ ).

#### 3.2. Discrete choice experiment to determine effect of product attributes on product choice

The intra-rater consistency determined by the weighted Kappa statistic ( $\kappa = 0.88$ ,  $p < 0.001$ ) indicated excellent agreement between product ratings in the two fixed tasks signifying high intra-participant reliability. Protein source (Utility score; mean  $\pm$  SD:  $137 \pm 70$ ;  $34 \pm 18\%$ ) and health ( $115 \pm 63$ ,  $29 \pm 16\%$ ) were the most important drivers of product choice, followed by sustainability ( $78 \pm 39$ ,  $19 \pm 10\%$ ), and origin ( $70 \pm 49$ ;  $9 \pm 6\%$ ).

##### 3.2.1. Plackett-Luce modelling to determine product worth of concepts included in the DCE

Final ranking of all 24 product concepts evaluated in the DCE, using Plackett-Luce rankings, is presented in Fig. 2. Soy product packaging with a 1.5 health star rating, imported ingredients, and absence of a sustainability tick was ranked lowest. Beef with a 4.5 health-star rating, made from NZ ingredients, and with a sustainability tick was ranked highest. A model only including main effects was selected to determine effects of attributes on product worth, because the AIC of the model did not reduce markedly by allowing for interaction effects, and the main effects model R-squared ( $R^2$ ) was so high it was deemed unsurpassable ( $R^2 = 0.96$ ). Therefore, interactions between attributes were deemed not relevant. A linear model on the Plackett-Luce rankings determined that product worth (as determined by the coefficient estimate of the model) was higher for beef than for soy (CE beef vs. soy:  $1.22 \pm 0.10$ ;  $p < 0.001$ ) and chickpeas (CE beef vs. chickpeas:  $0.55 \pm 0.10$ ;  $p < 0.001$ ), and higher for chickpeas than for soy (CE chickpeas vs. soy:  $0.66 \pm 0.10$ ;  $p < 0.001$ ). Furthermore, 4.5 health stars (CE 4.5 vs. 1.5 health stars:  $0.7 \pm 0.04$ ;  $p < 0.001$ ), NZ origin of ingredients (CE NZ vs. imported ingredients:  $0.45 \pm 0.04$ ;  $p < 0.001$ ), and the presence of a sustainability tick (CE present vs. absent sustainability tick:  $0.45 \pm 0.04$ ;  $p < 0.001$ ) were favoured, respectively, over 1.5 health star rating, the absence of a sustainability tick and imported ingredients.

##### 3.2.2. Differences in importance between product attributes

Importance of attributes, as assessed by mean square of the linear

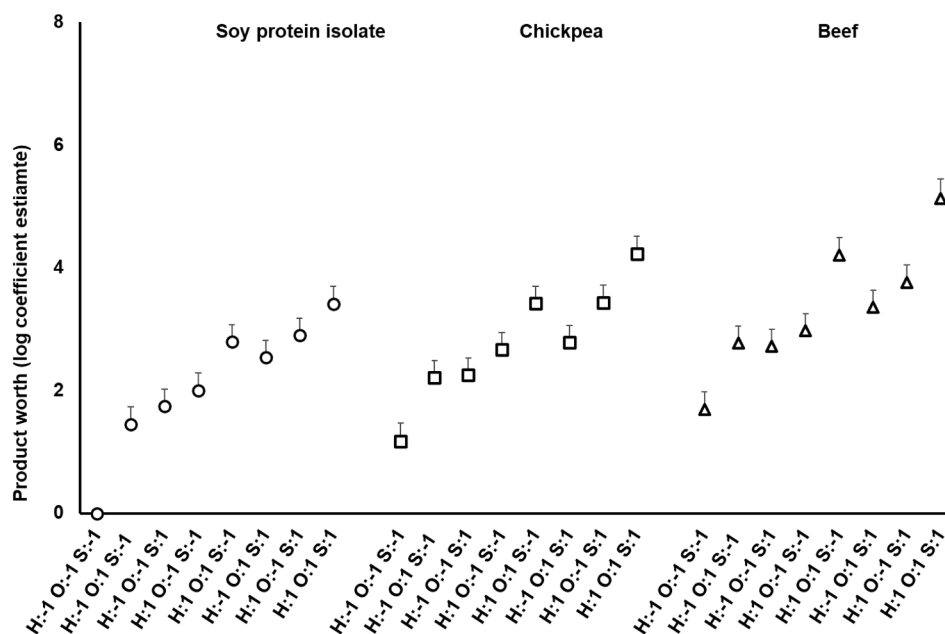


Fig. 2. Product worths (mean  $\pm$  SD) of each product ( $n = 24$ ) in the discrete choice experiment. Product worths are determined by Plackett-Luce modelling using the log coefficient estimates for each individual product. ‘1’ and ‘-1’ indicates whether the health (H), Origin (O) and Sustainability (S) attributes are present in its ‘high’ (i.e. 4.5 health stars, made from NZ ingredients, sustainability tick present) or ‘low’ (i.e. 1.5 health stars, made from imported ingredients, sustainability tick absent) attribute level. ○: product worth of soy protein isolate products; □: product worth of chickpea products; △: product worth of beef products.

model (Table 5), was highest for health (MS = 10.94), followed by protein source (MS = 5.57) origin of ingredients (MS = 5.02) and sustainability (MS = 4.93). When excluding each protein from the model in turn, the relative importance of health, origin and sustainability did not change markedly in terms of variance (Table 5). However, relative importance of protein source changed depending on which proteins were included in the analyses. When beef was compared to soy, the protein source's relative importance remained relatively high (MS = 5.96), whereas its importance became negligible when comparing beef with chickpea (MS = 1.23) or chickpea with soy (MS = 1.77). This highlights that the protein attribute was important in the decision between beef and soy, but markedly less important for other protein comparisons. Furthermore, the models indicate that the health attribute was the most consistent driver for product choice, whereas the protein attribute as a driver was more dependent on which options were included for the participant to choose between. Lastly, origin and sustainability were less important than protein source when beef was compared to a soy product; however, they became relatively more important than the protein attribute when any other comparison between proteins was made.

### 3.3. Participant segmentation based on utility ranges from the DCE

Elbow and Silhouette methods indicated that segmentation into two clusters was optimal based on utility ranges from the DCE. When comparing the clusters, the subsequently named 'Protein' Cluster (n = 67) had significantly higher utility ranges for protein source (196 ± 4; Table 6) than the nominated 'Health/Sustainability' Cluster (n = 77; utility score for protein source: 87 ± 3; p < 0.001). In contrast, the Health/Sustainability Cluster had higher utility scores health (157 ± 4; p < 0.001) and sustainability (87 ± 3; p < 0.001). Utility ranges for ingredient origin were not different between clusters. In terms of demographics, clusters had similar ( $\chi^2 = 0.10$ ; p = 0.75) gender ratios (Protein Cluster: 19 men, 48 women; Health/Sustainability Cluster: 20 men, 57 women), but the Protein Cluster had more ( $\chi^2 = 9.1$ ; p = 0.003) omnivores (n = 43, 64%) than flexitarians (n = 24, 36%), whereas the Health/Sustainability Cluster had more flexitarians (n = 47, 61%) than omnivores (n = 30, 39%).

### 3.4. Sample evaluation phase

#### 3.4.1. Effect of protein source on blind (Set 1) compared to aware (Set 2) sample tasting

Presentation order affected liking ratings in blind (Set 1) sample tasting. Samples rated third were rated lower than those that were rated first (p = 0.016) or second (p < 0.001), highlighting the importance of the balanced sample presentation.

A protein-source effect (p < 0.001) was evident on blind sample tasting and informed liking ratings, with higher ratings for beef (blind

**Table 5**

Mean squares of product attributes for the full linear model, and models for each of the protein pairs.

	Mean squares All proteins included	Beef vs. soy	Beef vs. chickpea	Chickpea vs. soy
Protein source	5.97	5.96	1.23	1.77
Health star rating	10.94	7.40	7.32	7.14
Ingredient origin	5.02	3.88	3.43	2.78
Sustainability tick	4.93	3.67	2.67	3.58

Mean squares for the protein, health, origin, and sustainability product attributes, of a linear model predicting product worth of each of the 24 products included in the DCE. Model allowed for main effects only.

**Table 6**

Participant clusters based on range of attribute utility scores.

		Protein Cluster (n = 67)	Health/ Sustainability Cluster (n = 77)
Gender	Women	48 (72%)	57 (74%)
	Men	19 (28%)	20 (26%)
Diet <sup>a</sup>	Omnivore	43 (64%)	30 (39%)
	Flexitarian	24 (36%)	47 (61%)
Utility score range	Protein source <sup>b</sup>	196 ± 4	87 ± 3
	Health star rating <sup>b</sup>	65 ± 3	157 ± 4
	Origin of ingredients	71 ± 4	70 ± 4
	Sustainability tick <sup>b</sup>	67 ± 3	87 ± 3

<sup>a</sup> indicates a difference in diet ratios between clusters, as determined with Chi-square (Chi: 9.1; p = 0.003).

<sup>b</sup> Indicates a difference in attribute utility scores between clusters (unpaired t-tests (p < 0.001)).

sample tasting: 81.3 ± 0.8) than soy (61.9 ± 1.3; p < 0.001) and chickpea (61.1 ± 1.7; p < 0.001) samples. Informing the participants of the protein source (Set 2) slightly reduced Set 1 *actual* liking ratings of beef (p = 0.005) and soy (p = 0.038) samples, but not chickpea (p = 0.40). *Actual* liking ratings exceeded *expected* liking ratings for beef (p < 0.001) and soy (p < 0.001), but were similar for chickpea (p = 0.08) when participants were aware (Set 2) of the protein source (Fig. 3).

#### 3.4.2. Effects of health star rating on expected and actual liking ratings

For beef samples, a 4.5-star health star rating increased *expected* liking ratings compared to the absence of the attribute (main effect of health: p < 0.001; Table 7). In contrast, in the chickpea sample, a 1.5-health star rating reduced *expected* liking ratings compared to the absence of the health attribute (p < 0.001). *Actual* liking ratings of the 4.5-star sample were higher than the 1.5-star product for all protein sources (all p ≤ 0.002).

#### 3.4.3. Effects of origin on expected and actual liking ratings

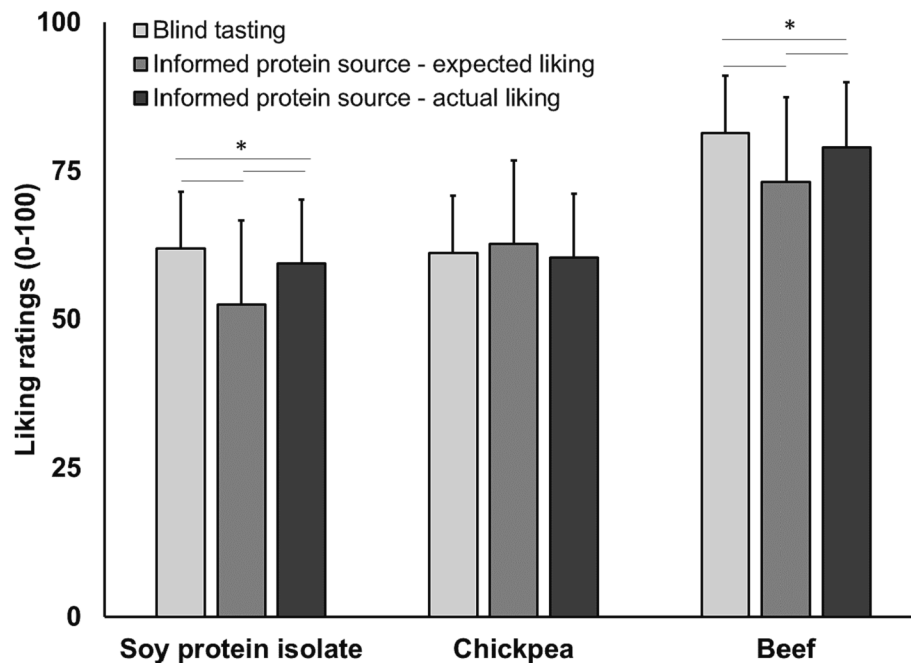
Samples made from NZ ingredients were *expected* to be liked better than products made with imported ingredients for all protein sources (main effect of origin: all p ≤ 0.05; Table 7). For beef samples, *actual* likings ratings were higher for those made with NZ ingredients compared to those made from imported ingredients (p < 0.001). In contrast, *actual* liking ratings of soy and chickpea samples were similar between samples made from imported and NZ ingredients (p > 0.05).

#### 3.4.4. Effects of sustainability tick on expected and actual liking ratings

Regardless of protein source, samples with a sustainability tick were *expected* to be liked more than products without a sustainability tick (main effect of sustainability: all p ≤ 0.007; Table 7). However, this effect was not detected on *actual* liking ratings for any protein source (p > 0.05). Beef products were liked more than expected regardless of the attribute level, but more so when the sustainability tick was absent than present (p = 0.005). For soy and chickpea samples, those without a sustainability tick were liked more, and samples with a sustainability tick were liked less, than expected (p < 0.001). Accordingly, *actual* liking of products with absent and present sustainability ticks gravitated towards a similar mean such that, in contrast to *expected* liking, the sensory experience was not different between products with different sustainability attribute levels.

#### 3.4.5. Interactions

For soy samples, interaction effects of health by sustainability (p = 0.025) and origin by sustainability (p = 0.030) on *expected* liking ratings existed. Soy products with 1.5 health stars and without a sustainability



**Fig. 3.** Ratings of expected and actual liking (mean ± SD) for soy protein isolate, chickpea, and beef samples during blind tasting (Sample set 1) and tasting with information of the protein source (Sample set 2). \* Vertical lines indicate significant differences between testing conditions ( $p < 0.05$ ). Paired t-tests were used to compare expected and actual liking ratings between Sample set 1 and Sample set 2 for each separate protein source.

**Table 7**

Main effects of protein source, health star rating, origin of ingredients and presence of a sustainability tick on expected, actual and the difference between expected and actual liking ratings.

Attribute	Protein source	Attribute level	Expected liking rating (Mean ± SD)	Main effect (P value)	Actual liking rating (Mean ± SD)	Main effect (P value)	Difference between expected and actual liking (Mean ± SD)	Main effect (P value)
Health star rating	Beef	Attribute absent	73.8 ± 13.7 <sup>a</sup>	<0.001	78.9 ± 11.2 <sup>a,b</sup>	<0.001	5.1 ± 11.8	0.18
		1.5 health stars	70.8 ± 15.5 <sup>a</sup>		76.1 ± 13.1 <sup>a</sup>		5.2 ± 11.2	
		4.5 health stars	78.2 ± 12.7 <sup>b</sup>		81.4 ± 12.1 <sup>b</sup>		3.2 ± 11.4	
	Soy protein isolate	Attribute absent	53.8 ± 15 <sup>a,b</sup>	<0.001	58.4 ± 15.5 <sup>a,b</sup>	<0.001	4.7 ± 15.5 <sup>a</sup>	<0.001
		1.5 health stars	53.4 ± 19.3 <sup>a</sup>		54.3 ± 19.1 <sup>a</sup>		0.9 ± 15.3 <sup>a</sup>	
		4.5 health stars	59.3 ± 16.9 <sup>b</sup>		59.4 ± 16.6 <sup>b</sup>		0.2 ± 15.1 <sup>a</sup>	
Chickpea	Attribute absent	63.1 ± 15.1 <sup>a</sup>	<0.001	61.3 ± 19.4 <sup>a,b</sup>	0.002	-1.7 ± 15.7	0.09	
	1.5 health stars	56.6 ± 18.8 <sup>b</sup>		56.0 ± 22.5 <sup>a</sup>		-0.6 ± 12.7		
	4.5 health stars	63.8 ± 19.2 <sup>a</sup>		60.4 ± 22.6 <sup>b</sup>		-3.4 ± 13.0		
Origin of ingredients	Beef	Attribute absent	73.7 ± 14.4 <sup>a,b</sup>	<0.001	78.7 ± 11.7 <sup>a</sup>	<0.001	5.0 ± 11.6	0.61
		Imported ingredients	71.5 ± 14.1 <sup>a</sup>		76.4 ± 12.5 <sup>a</sup>		4.9 ± 12.4	
		NZ ingredients	78.0 ± 12.7 <sup>b</sup>		81.7 ± 11.5 <sup>b</sup>		3.7 ± 10.9	
	Soy protein isolate	Attribute absent	55.5 ± 15.8 <sup>b</sup>	0.05	59.9 ± 16.1 <sup>a</sup>	0.003	4.3 ± 15.6	0.10
		Imported ingredients	52.2 ± 18.9 <sup>a</sup>		54.1 ± 16.8 <sup>b</sup>		1.9 ± 15.5	
		NZ ingredients	56.2 ± 16.7 <sup>b</sup>		56.4 ± 17.3 <sup>a,b</sup>		0.2 ± 14.8	
Chickpea	Attribute absent	61.1 ± 16.3 <sup>a,b</sup>	0.03	60.1 ± 20.4	0.54	-1.0 ± 14.9	0.18	
	Imported ingredients	60.8 ± 18.4 <sup>a</sup>		59.1 ± 21.8		-1.7 ± 12.5		
	NZ ingredients	64.5 ± 17.8 <sup>b</sup>		60.4 ± 21.6		-4.2 ± 15.4		
Sustainability tick	Beef	Attribute absent	73.2 ± 14.4 <sup>a,b</sup>	<0.001	78.1 ± 12.0	0.28	4.9 ± 12.2 <sup>a,b</sup>	0.005
		Tick absent	72.2 ± 14.5 <sup>a</sup>		79.0 ± 12.2		6.8 ± 10.8 <sup>b</sup>	
		Tick present	78.6 ± 12.0 <sup>b</sup>		80.8 ± 11.6		2.2 ± 10.3 <sup>a</sup>	
	Soy	Attribute absent	54.7 ± 17.1 <sup>a,b</sup>	<0.001	59.1 ± 17.3	0.86	4.3 ± 15.4 <sup>a,b</sup>	<0.001
		Tick absent	51.9 ± 16.9 <sup>a</sup>		56.0 ± 15.6		4.1 ± 16.1 <sup>a</sup>	
		Tick present	58.4 ± 15.4 <sup>b</sup>		56.1 ± 16.0		-2.3 ± 14.1 <sup>b</sup>	
Chickpea	Attribute absent	60.2 ± 16.9 <sup>a</sup>	0.007	57.9 ± 21.2	0.23	-2.3 ± 14.2 <sup>a</sup>	<0.001	
	Tick absent	61.6 ± 16.9 <sup>a</sup>		63.1 ± 20.4		1.5 ± 12.5 <sup>b</sup>		
	Tick present	65.9 ± 17.6 <sup>b</sup>		61.8 ± 20.5		-4.0 ± 16.6 <sup>a</sup>		
Interaction effects	Beef	Health*origin*sustainability		NS		NS	0.04	
		Health*sustainability		0.025		NS	NS	
		Origin*sustainability		0.003		NS	NS	
	Chickpea	Health*		NS		NS	0.04	

Main effects were determined by using mixed effects models, with participants as a random effect, and gender, age, diet, meat attachment score, cluster adherence, protein source, health, origin of ingredients and sustainability as fixed factors. <sup>a,b</sup> different letters indicate differences between attributes levels within an attribute according to estimated marginal means with Benjamini Hochberg corrections for multiple comparisons within attribute levels for a protein source.

tick were expected to be liked less than products without a health and sustainability attribute ( $p = 0.03$ ) or products with 4.5 health stars and a sustainability tick ( $p < 0.001$ ). Soy products made from imported ingredients and without a sustainability tick were expected to be liked less than products without an origin or sustainability attribute ( $p = 0.005$ ), products that only had an origin or sustainability attribute (regardless of the attribute level; all  $p < 0.05$ ), and products that were made from NZ ingredients and exhibited a sustainability tick ( $p < 0.001$ ). Interaction effects indicated for beef and chickpea were not detected in post hoc effects ( $p > 0.05$ ).

#### 3.4.6. Effects of gender and cluster membership

Gender, diet, meat attachment scores and cluster membership had no impact on either blind or informed sample liking scores ( $p > 0.05$ ). In the Health/Sustainability compared to the Protein Cluster, *expected* liking ratings were higher for soy samples (Protein Cluster:  $51.8 \pm 17.2$ ; Health/Sustainability Cluster:  $57.6 \pm 15.9$ ; main effect of cluster  $p = 0.006$ ), and *actual* liking ratings were higher for chickpea samples (Protein Cluster:  $56.3 \pm 22.0$ ; Health/Sustainability Cluster:  $63.1 \pm 19.5$ ;  $p = 0.04$ ).

#### 3.4.7. Actual liking ratings are predicted by expected liking ratings

R-squared ( $R^2$ ) values of four different models predicting actual liking ratings were compared. The null-model, not including any fixed factors (participant as a random effect), had the lowest  $R^2$  value (0.21). The  $R^2$  value of the model only with expected liking ratings ( $R^2 = 0.56$ ) was higher than the model with only product attributes (protein, health, origin, sustainability;  $R^2 = 0.45$ ), indicating that expected liking ratings have a larger effect on actual liking ratings than the effect of the product attributes themselves. The robustness of the model only increased marginally ( $R^2 = 0.62$ ) when the full model was compared to the model with only expected liking ratings (all comparisons between models  $p < 0.001$ ).

## 4. Discussion

This study aimed to determine the effects of protein source, health, origin, and sustainability attributes on perception of a product at point of purchase using a DCE, as well as how such attributes affected sample perception during sensory evaluation. Protein and health attributes had the larger and more consistent effects on product choice and sensory experience than origin and sustainability attributes. Furthermore, the finding that there are two clear consumer groups - one driven by protein source regardless of health and environment attributes, and one driven by health and to a lesser extent sustainability irrespective of protein source - is noteworthy.

#### 4.1. Protein source is a main driver for product choice and sensory evaluation

Protein source was, alongside health, the most important driver of product choice in the DCE, and the most decisive driver of *expected* and *actual* liking ratings during sensory evaluation. Interestingly, consumers had a strong preference for beef when the choice was between beef and soy, but protein source was less important when the choice was between beef and chickpeas, or soy and chickpeas. This highlights that the variability in protein sources included in this study caused much of the variation in importance, and it leaves the questions whether effect of protein source may have been less important if only plant-based sources were included. Consistent with this study, previous DCEs including meat and PBMA have reported that beef/meat products are preferred over PBMA in consumers who consume meat in their diet (Asioli et al. 2022; Caputo et al. 2022; Profeta et al. 2021; Slade 2018; Van Loo et al. 2020), although vegetarians are more likely to prefer PBMA (Slade 2018; Van Loo et al. 2020). Therefore our findings are limited to a consumer group that does not eliminate meat from their diet. The finding that soy is

negatively perceived agrees with previous studies where the presence of the word 'soy' on a product package negatively affected sensory evaluation of a nutrition bar. Nutrition bars were perceived to be more grainy and less flavourful when the word 'soy' was present on the label, compared to when it was absent (Wansink et al. 2000). Furthermore, in an online survey among 534 Swiss consumers, a chicken alternative made from pea protein was perceived to be less natural (rank 14 out of 20 products) than chickpeas (rank 2/20), beef entrecote (rank 5/20), or falafel (rank 10/20) (Hartmann et al. 2022) – perhaps indicating that a product from soy protein may similarly be perceived as inferiorly natural to a beef or chickpea product in the current study, and play a large role in product choice through perceived naturalness (Román et al., 2017).

#### 4.2. A high health star rating influences product choice and sensory experience

The health attribute was identified as a main driver of product choice (DCE), consistent across all protein sources, with preference for products with 4.5, compared to 1.5, health stars. The health attribute was also the only attribute, apart from protein source, that consistently affected *actual* liking ratings when coupled with the sensory experience. In previous literature, consumers also prioritised health motivators over other motivators. For example, in an online survey (Lang 2020), US consumers indicated that their top three reasons for consuming hybrid meat/vegetable products instead of full meat products were health related, ranked above reasons of price, taste, and sustainability. Although it has been reported previously that health is a major driver for PBMA consumption in general (Onwezen et al. 2021), the data from the current study identifies further that this is the case irrespective of the protein source of the product.

In contrast to the current study, previous studies have often not found effects of health claims (e.g. 'low in salt', 'light', 'low fat', 'high protein') on *actual* liking during sensory evaluation, despite their effects on *expected* liking ratings (Kähkönen and Tuorila 1998; Liem et al. 2012; Norton et al. 2013; Rramani et al. 2023). It needs to be emphasised that the effects found on liking ratings in the current study were reliant on the presence of a negatively perceived health attribute, since differences between the absence of the health attribute and a 4.5 health star rating were found only for *expected* liking ratings of beef samples, and were not present for *actual* liking ratings of any protein source. Therefore, the addition of a positive health claim as opposed to no health claim may have minimal effects on product choice of PBMA in practice. However, consumers more positively experienced a highly health-rated PBMA than a low health-rated PBMA, possibly because health was a major driver of why they selected a PBMA in the first place.

#### 4.3. Claims relating to origin of ingredients and environmental sustainability influence product choice and expected liking, but not sensory experience

Importance of origin and sustainability attributes during the product choice and sensory experience was relatively low compared to the protein source and health attributes. Previously, additional information on health and/or environment did not alter liking of a negatively experienced faba bean spread, whereas health and environment information increased liking of better-liked chickpea and black bean spreads (Henn et al. 2023). Similarly, this study showed that the inclusion of a positively perceived origin message ('made from NZ ingredients') improved liking only in the most-liked sample (i.e. beef), but did not affect liking of less liked soy and chickpea samples. In addition, the 'made from NZ ingredients' label did not significantly increase *expected* liking ratings compared to the absence of the attribute, and the presence of a sustainability tick increased *expected* liking ratings in chickpea samples only. This may indicate that consumers mostly ignored the origin message and sustainability tick in their sample assessment and/or

that consumers did not believe that consuming PBMA instead of meat contributes to environmental sustainability. In support of the latter, some studies report that reducing meat consumption was generally perceived by consumers to have a low impact on global warming, and their willingness to change meat consumption, compared to other initiatives, to mitigate global warming was low (de Boer et al. 2016; Truelove and Parks 2012). Alternatively, consumers may recognise that PBMA are more sustainable than meat, but due to lack of strong ethical value orientation, still prefer meat over PBMA (Hoek et al. 2011).

#### 4.4. A 'Protein Cluster' and a 'Health/Sustainability Cluster' were established

In the current study, a Protein and a Health/Sustainability Cluster were established based on participants' ranges of utility scores of the product attributes. It has been indicated previously that there is an association between those who have more positive perceptions of plant-based diets and those who consume meat less frequently: an online survey of Danish consumers (Reipurth et al. 2019) established that, low, compared to high, animal product consumers, perceived plant-based diets/products to be healthier and better for environmental sustainability (Reipurth et al. 2019). In this current study, the higher proportion of omnivores in the Protein Cluster than in the Health/Sustainability Cluster may indicate a reduced belief in the value or accuracy of health and environment-related claims on PBMA, or higher belief in these claims with regards to meat, hence favouring beef over PBMA. Although the health, origin, and sustainability attributes were not perceived differently between clusters during the sensory experience, the finding that the Health/Sustainability Cluster view health and sustainability as more important, and protein source as less important drivers for product choice, than the Protein Cluster, highlights that about half of consumers could be influenced by extrinsic health and sustainability attributes in their decision-making between meat- and plant-based products.

#### 4.5. Expected compared to actual liking ratings during sensory evaluation

An important finding is that concept evaluation during the sensory experience correlated with *expected* liking ratings for the product concept before the sensory experience. This infers that, in general, the sensory experience lived up to expectations evoked by the product packaging, as indicated by similar *expected* and *actual* liking ratings. These findings are in agreement with previous research reporting that consumer evaluations shift towards their previous expectations as long as any contrast between expectations and reality is relatively small, to minimise the difference between the two and avoid cognitive dissonance (Piqueras-Fiszman and Spence 2015; Schifferstein et al. 1999).

An exception to this finding was that, despite the expected effects of sustainability attribute levels on *expected* liking, the sensory experience was not different between products with different sustainability attribute levels, regardless of the protein source. In other words, although the presence of a sustainability tick was expected to have a positive effect on sample evaluation, this effect was not experienced in reality.

## 5. Limitations

The strengths of this study include the use of a DCE in combination with sample evaluation, so that attributes effects could be determined at point of purchase, as well as during product experience. Furthermore, the fractional design enabled the effect of product attributes alone, as well as in combination with each other to be determined. However, a limitation of this study is that it did not measure how the different product attributes were interpreted by the consumers. Effects of the attributes on perceived naturalness, healthiness and environmental sustainability would be of interest. Secondly, in reality, a main ingredient of soy protein isolate may not be as prominently presented on the

front-of-package as it was in this study. Therefore, negative evaluation of the soy products in this study may have been overestimated. Furthermore, during the sample evaluation phase, absence of a sustainability tick looked identical to the absence of the sustainability attribute altogether. Therefore, a product concept with the absence of a tick may not necessarily have been interpreted as an unsustainable product. This may have resulted in understating the importance of sustainability. Last, familiarity with the health star ratings and sustainability tick and believability of the package information was not measured, but could have affected the evaluation of the products (Ang et al. 2023; van Herpen et al. 2012).

## 6. Conclusions

This study established that protein source, health star rating, origin of ingredients, and the presence of a sustainability tick all affect product choice and *expected* liking ratings. During sensory evaluation, a highly health-rated product was experienced better than a product with a low health rating, but the effects of origin and sustainability attributes on sample experience were minimal for the plant-based samples. Not all consumers were equally sensitive to the presence of extrinsic product attributes: whereas there was one group of consumers who always preferred beef over soy and chickpea products, there was a substantial proportion of consumers that could be swayed by product attributes on front-of-pack labelling, making these consumers the most likely to engage with PBMA.

#### Ethical statement

The study was considered and assessed as low risk following the Massey University Human Ethics Committee process (human ethics notification number: 4000025647). All respondents indicated informed consent by checking a checkbox to that effect at commencement of the survey.

#### Funding

This work was funded by a NZ Ministry of Business and Education (MBIE) Catalyst Grant (MAUX2001).

#### CRediT authorship contribution statement

**Caroline Giezenaar:** Conceptualization, Methodology, Formal analysis, Investigation, Project administration, Writing – original draft, Writing – review & editing. **A. Jonathan R. Godfrey:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Meika Foster:** Funding acquisition, Supervision, Writing – review & editing. **Joanne Hort:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

#### 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

Data will be made available on request.

#### Acknowledgements

The authors wish to acknowledge the support of Dianne Fountaine-Cody, Karina Gonzalez Estanol, and Rebekah Orr for their assistance with the study sessions.

## References

- Ang, M. Y., Pontes, N., & France, C. (2023). The influence of health star rating labels on plant-based foods: the moderating role of consumers' believability. *Food Quality and Preference*, 107, Article 104827.
- Asioli, D., et al. (2022). European consumers' valuation for hybrid meat: does information matter? *Applied Economic Perspectives and Policy*. n/a (n/a).
- Caputo, V., Sogari, G., Loo, V., & Ellen, J. (2022). Do plant-based and blend meat alternatives taste like meat? A combined sensory and choice experiment study. *Applied Economic Perspectives and Policy*. n/a (n/a).
- Collier, E. S., et al. (2021). Identifying barriers to decreasing meat consumption and increasing acceptance of meat substitutes among Swedish consumers. *Appetite*, 167, Article 105643.
- Cozy Corner 'Realistic Kitchen Ambience ASMR | Kitchen Sounds and Animations', <[https://www.youtube.com/watch?v=IA\\_Q5FlzxmK](https://www.youtube.com/watch?v=IA_Q5FlzxmK)>, accessed 2020.
- de Boer, J., de Witt, A., & Aiking, H. (2016). Help the climate, change your diet: a cross-sectional study on how to involve consumers in a transition to a low-carbon society. *Appetite*, 98, 19–27.
- Elzerman Johanna, E., van Boekel Martinus, A. J. S., & Luning Pieter, A. (2013). Exploring meat substitutes: consumer experiences and contextual factors. *British Food Journal*, 115(5), 700–710.
- Giezenaar, C., & Hort, J. (2021). A narrative review of the impact of digital immersive technology on affective and sensory responses during product testing in digital eating contexts. *Food Research International*, 150, Article 110804.
- Good Food Institute (2022), '2021 U.S. retail market data for the plant-based industry'.
- Graça, J., Oliveira, A., & Calheiros, M. M. (2015). Meat, beyond the plate data-driven hypotheses for understanding consumer willingness to adopt a more plant-based diet. *Appetite*, 90, 80–90.
- Hartmann, C., Furtwaengler, P., & Siegrist, M. (2022). Consumers' evaluation of the environmental friendliness, healthiness and naturalness of meat, meat substitutes, and other protein-rich foods. *Food Quality and Preference*, 97, Article 104486.
- Heller, Martin C and Keoleian, Gregory A (2018), 'Beyond meat's beyond burger life cycle assessment: a detailed comparison between'.
- Henn, K., et al. (2023). Health versus environmental benefits: Does additional information influence consumer acceptance of pulse-based spreads? *Journal of Food Science*, 88(3), 1144–1158.
- Hoek, A. C., et al. (2011). Replacement of meat by meat substitutes. A survey on person- and product-related factors in consumer acceptance. *Appetite*, 56(3), 662–673.
- Kähkönen, Päivi, & Tuorila, H. (1998). Effect of reduced-fat information on expected and actual hedonic and sensory ratings of sausage. *Appetite*, 30(1), 13–23.
- Kassambara, A. and Mundt, F. (2020), 'factoextra: Extract and Visualize the Results of Multivariate Data Analyses'.
- Kerslake, E., Kemper, J. A., & Conroy, D. (2022). What's your beef with meat substitutes? Exploring barriers and facilitators for meat substitutes in omnivores, vegetarians, and vegans. *Appetite*, 170, Article 105864.
- Lang, M. (2020). Consumer acceptance of blending plant-based ingredients into traditional meat-based foods: evidence from the meat-mushroom blend. *Food Quality and Preference*, 79, Article 103758.
- Lazzarini, G. A., et al. (2016). 'Does environmental friendliness equal healthiness? Swiss consumers' perception of protein products'. *Appetite*, 105, 663–673.
- Liem, D. G., Toraman Aydin, N., & Zandstra, E. H. (2012). Effects of health labels on expected and actual taste perception of soup. *Food Quality and Preference*, 25(2), 192–197.
- Liu, R., et al. (2017). A natural experiment: using immersive technologies to study the impact of "all-natural" labeling on perceived food quality, nutritional content, and liking. *Journal of Food Science*, 82(3), 825–833.
- Macdiarmid, J. I. (2022). The food system and climate change: are plant-based diets becoming unhealthy and less environmentally sustainable? *Proceedings of the Nutrition Society*, 81(2), 162–167.
- Meyer, D., Zeileis, A., and Hornik, K. (2023), 'vcd: Visualizing Categorical Data'.
- Michel, F., Hartmann, C., & Siegrist, M. (2021a). Consumers' associations, perceptions and acceptance of meat and plant-based meat alternatives. *Food Quality and Preference*, 87, Article 104063.
- Michel, F., et al. (2021b). A multi-national comparison of meat eaters' attitudes and expectations for burgers containing beef, pea or algae protein. *Food Quality and Preference*, 91, Article 104195.
- Norton, J. E., Fryer, P. J., & Parkinson, J. A. (2013). The effect of reduced-fat labelling on chocolate expectations. *Food Quality and Preference*, 28(1), 101–115.
- Omre, B. (2014). *Including holdout choice tasks in conjoint studies*. Sawtooth Software Research Paper Series.
- Onwezen, M. C., et al. (2021). A systematic review on consumer acceptance of alternative proteins: pulses, algae, insects, plant-based meat alternatives, and cultured meat. *Appetite*, 159, Article 105058.
- Piqueras-Fiszman, B., & Spence, C. (2015). Sensory expectations based on product-extrinsic food cues: an interdisciplinary review of the empirical evidence and theoretical accounts. *Food Quality and Preference*, 40, 165–179.
- Profeta, A., et al. (2021). Discrete choice analysis of consumer preferences for meathybrids-findings from Germany and Belgium. *Foods*, 10(1), 17.
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Reipurth, M. F. S., et al. (2019). Barriers and facilitators towards adopting a more plant-based diet in a sample of Danish consumers. *Food Quality and Preference*, 73, 288–292.
- Román, S., Sánchez-Siles, L. M., & Siegrist, M. (2017). The importance of food naturalness for consumers: results of a systematic review. *Trends in Food Science & Technology*, 67, 44–57.
- Rramani, Q., et al. (2023). Nutrition claims influence expectations about food attributes, attenuate activity in reward-associated brain regions during tasting, but do not impact pleasantness. *Brain and Behavior*, 13(1), Article e2828.
- Sawtooth Software Inc. (2021), 'The CBC/HB System for Hierarchical Bayes Estimation', *Sawtooth Software Technical Paper Series*.
- Schiffenstein, H. N. J., Kole, A. P. W., & Mojet, J. (1999). Asymmetry in the disconfirmation of expectations for natural yogurt. *Appetite*, 32(3), 307–329.
- Schutz, H. G., & Cardello, A. V. (2001). A labeled affective magnitude (lam) scale for assessing food liking/disliking. *Journal of Sensory Studies*, 16(2), 117–159.
- Siegrist, M., & Hartmann, C. (2019). Impact of sustainability perception on consumption of organic meat and meat substitutes. *Appetite*, 132, 196–202.
- Slade, P. (2018). 'If you build it, will they eat it? Consumer preferences for plant-based and cultured meat burgers'. *Appetite*, 125, 428–437.
- Truelove, H. B., & Parks, C. (2012). Perceptions of behaviors that cause and mitigate global warming and intentions to perform these behaviors. *Journal of Environmental Psychology*, 32(3), 246–259.
- Tso, R., Lim, A. J., & Forde, C. G. (2021). A critical appraisal of the evidence supporting consumer motivations for alternative proteins. *Foods*, 10(1), 24.
- Turner, H. L., et al. (2020). Modelling rankings in R: the PlackettLuce package. *Computational Statistics*, 35(3), 1027–1057.
- van Herpen, E., Seiss, E., & van Trijp, H. C. M. (2012). The role of familiarity in front-of-pack label evaluation and use: a comparison between the United Kingdom and The Netherlands. *Food Quality and Preference*, 26(1), 22–34.
- Loo, V., Ellen, J., Caputo, V., & Lusk, J. L. (2020). Consumer preferences for farm-raised meat, lab-grown meat, and plant-based meat alternatives: Does information or brand matter? *Food Policy*, 95, Article 101931.
- VIABEL 'Supermarket SOUND europe RELAX, STUDY & Enjoy ASMR Ambient noise', <<https://www.youtube.com/watch?v=Dp9yBndP-1A&t=1199s>>, accessed 2019.
- Wansink, B., et al. (2000). How soy labeling influences preference and taste. *The International Food and Agribusiness Management Review*, 3(1), 85–94.
- Wickham, H., et al. (2022), 'dplyr: A Grammar of Data Manipulation'.