

Insights from interval-valued ratings of consumer products—a DECSYS appraisal.

Zack Ellerby, Oliver Miles, Josie McCulloch (Member, IEEE), Christian Wagner (Senior Member, IEEE)
LUCID Research Group, University of Nottingham, UK

zack.ellerby; oliver.miles; josie.mcculloch; christian.wagner@nottingham.ac.uk

Abstract—The capture and analysis of interval-valued data has seen increased interest over recent years. This offers a direct means to capture and reason about uncertainty in data, whether obtained from sensors or from people. Open-source software (DECSYS [1]) was recently released to facilitate the efficient capture of interval-valued survey responses. Potential real-world applications are broad ranging, and this paper documents an initial test-case of the software and its underpinning methodology, in a marketing-centric application. It provides an illustration of the insights offered by interval-valued responses, in this case relating to consumer preferences. We apply two approaches to describe and draw insights from the data: inferential statistics and descriptive visualisation methods. Statistical results indicate that overall purchase intention was well-described by four factors: value, healthiness, taste and brand. The capture of uncertainty information, afforded by intervals, also permitted identification of six factors that contribute to purchase intention uncertainty—relating to taste, ethics and visual appearance. Visualisations of interval-valued responses, using the IAA [2]–[5], also highlighted factors with high degrees of uncertainty—in particular, product ethics. This information could prove valuable for retailers in determining how to focus future marketing campaigns. It may prove equally valuable for market regulators, by informing where to improve product labelling information. More generally, the case study provides an overview of capturing and analysing intervals, highlighting some of the challenges, but also the unique potential to gain additional insights not available using conventional, ‘crisp’, approaches.

I. INTRODUCTION AND BACKGROUND

Collecting interval-valued responses offers the potential to obtain greater information content than point responses. The variable dimension of width that is possessed by an interval-value informs the degree of certainty, variability or fuzziness associated with the response in question. This information is often overlooked by traditional discrete, or point, response modes (such as Likert-type [6], or Visual Analogue [7] scales), for which the forced specificity may give a false impression of certainty or precision that is not truly warranted—in fact, this may increase the noise present in the response.

Conventional methods of obtaining interval-valued responses, e.g. through soliciting two separate (minimum and maximum) endpoints, are substantially more effortful and time consuming than equivalent point-response modes, for both survey participants and administrators. In combination with relatively low awareness concerning the potential benefits of

interval-valued data, this has contributed to low levels of general adoption of interval-valued data capture across the wider research community.

Recently released open-source software—‘DECSYS’[1]—permits electronic capture of interval-valued responses, using an ellipse response format that aims to be quick, simple and intuitive to use, thus maximising response information capture efficiency. Therefore, a tool now exists to facilitate the efficient capture of interval-valued responses, which may prove valuable across a range of research areas. However, these have yet to become widely adopted.

Complementary research studies are ongoing to address socio-technical considerations—concerning the basic efficacy of this response mode to capture multiple sources of respondent uncertainty and response variability—and also to demonstrate the value of the resulting data in real-world contexts [8]. The present study was conducted to demonstrate both the value provided by interval-valued responses and the use of the DECSYS software’s ellipse response mode, to capture said intervals. To achieve this, we describe the implementation of this method in the context of one of many real-world applications for which the additional information captured by interval-valued responses could provide valuable insights. The area chosen is an important and widely studied area of market research—that of consumer preferences, for which the breadth of literature ranges from wine [9]–[14], to apples [15]–[20] to electric vehicles [21]–[26]. In the present case, we investigate consumer ratings in relation to a range of snack food products.

Understanding consumer attitudes, and the factors that influence these, is fundamental to this far-reaching area of research [27]. Moreover, existing literature already acknowledges the importance of understanding consumer conviction, or attitude certainty—both in terms of influencing future choice behaviour and in determining openness to persuasion versus resistance to future attitude change [28]–[31]. This may be a key consideration in informing which attributes it may be most effective to focus advertising material upon, when the aim is to encourage or reinforce positive attitudes towards a product. Identification of areas for which consumers are highly uncertain may also be of value to market regulators, whose objective may instead be to ensure that consumers are sufficiently well-informed to make sensible purchase decisions.

This study represents a test case, designed to examine consumer values, product perceptions, and also the uncertainty

surrounding these, by capturing this data in interval-valued format using the DECSYS software. We use these data to build statistical models, and also provide data visualisations, using the Interval Agreement Approach [2]–[5], to demonstrate insights that may be provided; both into consumer values and product perceptions, and into how these contribute towards overall product purchase intention.

Potential benefits of obtaining interval-valued responses, as identified a priori, include:

- Added predictive value, in terms of overall purchase intention, provided by uncertainty associated with factor ratings—as indicated by interval width.
- Added information regarding uncertainty surrounding overall purchase intention—as indicated by interval width—as well as allowing identification of the factors associated with this.
- Added information regarding consumer uncertainty, in relation to both self-reported values and product attribute perceptions.

In addition to consumer preference ratings, we collect participant ratings concerning their use of the ellipse response mode itself. These focus on perceived ease of use, unnecessary complexity, effective communication and overall liking. This allows a secondary assessment of whether respondents subjectively felt that the new response format offered additional value, and whether this was at the cost of increased workload.

In Section II we describe the participants, stimuli and procedure of the experimental study—including details concerning the processes of both data collection and analysis. In Section III we report the findings of both descriptive and inferential strands of analysis of the data obtained in the study. In Section IV we summarise and provide a discussion of the key findings and lessons learned.

II. METHOD

A. Study Participants

A total of forty participants were recruited to take part in the study, in an opportunity sample. The study was open to anyone who wished to take part, but as it was conducted on campus at the University of Nottingham, it can be assumed that a substantially greater proportion of participants were students or staff of an academic institution than would be expected in the general population. Ages ranged from 18 to 55 ($M=24.7$, $SD=8.0$).

B. Experimental Stimuli

A total of eight high-street convenience food products—i.e. ‘snacks’—were selected as stimuli. Each participant viewed each product and its associated packaging, as well as tasting a small portion of each product when prompted to do so. The eight snack-food products were chosen to represent diversity along six dimensions—corresponding to the six attribute ratings made for each product (Visual Appeal, Value for Money, Healthiness, Taste, Branding and Ethics). For example, three products were chocolate bars—each of these were therefore of relatively low nutritional value and likely to be rated highly

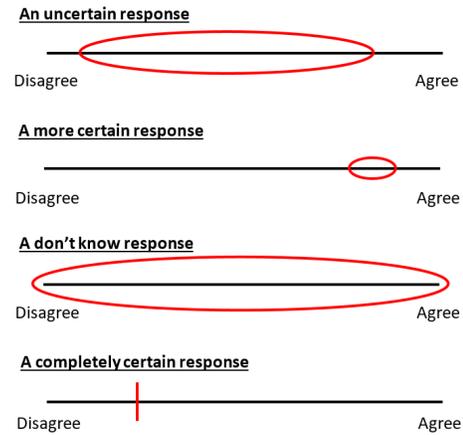


Fig. 1. Illustrative examples, as provided to study participants, showing interval-valued responses representing different degrees of uncertainty.

on taste—but one of these was a popular confectionery brand, another a budget supermarket own brand, and a third a brand that focuses on ethical sourcing at a higher price point. Other products are marketed as offering greater nutritional value, with options presented across a range of price points. The specific products sampled and rated in this study were¹:

- Tesco Raisin Munch Bar (supermarket own-brand)
- Tesco Milk Chocolate Bar (supermarket own-brand)
- Cadbury’s Dairy Milk Chocolate Bar
- Divine Fairtrade Milk Chocolate Bar
- Eat Natural Fruit & Nut Bar—Almond & Apricot
- Bounce Protein Energy Ball—Almond
- Nature Valley Crunchy Granola Bars—Oats & Honey
- Nākd Cocoa Delight Natural Vegan Snack Bar

C. Data Collection Procedure

The study procedure was approved by the Ethics Committee at the University of Nottingham School of Computer Science. Before beginning the task, participants were provided with two information sheets. One provided general information, concerning the nature of the study, the total estimated time to take part (30 minutes), the inconvenience allowance provided (£6.50), and their right to withdraw at any time. The second provided a brief explanation of the use of the ellipse response format, including basic example responses (cf. Fig. 1). Having had the opportunity to review these information sheets and ask questions, participants who wished to proceed signed the consent form and began the study.

The primary survey was both created and administered using the DECSYS software [1]—*Workshop Mode* was used, with all survey responses collected locally. Questions were presented via a Microsoft Surface Pro tablet computer. The touchscreen capabilities of this device enable users to interact with the screen via a stylus (Microsoft Surface Pen). This hardware is particularly well-suited to the DECSYS *ellipse* response mode, because it allows for precise and intuitive provision of ellipse

¹Please note that the aim of this study is not to provide evidence for the superiority of any particular product. We do not claim that our findings provide generalised evidence for real-world product quality or desirability.

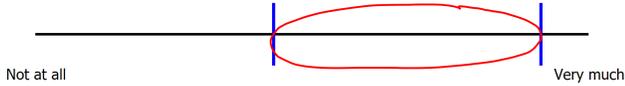


Fig. 2. Example survey response scale with authentic moderate uncertainty response, as drawn (using a digital pen) by a DECSYS user on a touchscreen display.

responses, with minimal differences from how this would be achieved using pen and paper.

Survey questions were administered in three key stages:

First, before product exposure, participants were asked to provide hypothetical value ratings concerning the importance of each of the six product attributes—e.g. ‘When making a product choice, how important is *visual appeal*?’

Second, during product exposure, participants were asked, as appropriate, either to view or taste each product, or consider its packaging or information provided on this, and to provide a rating for the associated product attribute. Specific questions for each attribute are shown in Table I.

Third, after product exposure, participants were asked to provide overall ratings for their purchase intention towards each product.

For each question, the response scale was continuous, without visual markers to indicate increments or specific intermediate values. Pen colour was red, and range markers blue—see Fig. 2 for illustration. The scale was labelled at the far left with ‘Not at all’ and at the far right with, ‘Very much’. The survey was programmed in such a way as to ask each of the six attribute questions for each product. Question order was randomised within each section of the survey (i.e. Before, During and After product exposure). This was to keep participants engaged in reading each question considerably before responding and to preclude any potential order effects.

Following completion of the primary product rating survey, participants were also asked to provide their level of agreement with four statements concerning their subjective user experience of the ellipse response format, as administered through the DECSYS software. These were adapted from the ‘Systems Usability Scale’ [32], and administered separately, on paper, and using a traditional 5-point ordinal response scale ranging from 1—‘Strongly Disagree’, to 5—‘Strongly Agree’. The questions asked were as follows:

- ‘I found the response-format easy to use’
- ‘I found the response-format unnecessarily complex.’
- ‘I found that the response-format allowed me to effectively communicate my desired response.’
- ‘Overall, I liked the response-format.’

Following completion of the study participants were paid an inconvenience allowance of £6.50, they were also offered the opportunity to spend up to £1.50 from this allowance to purchase products that they had sampled at a rate discounted from retail value.

D. Analysis Procedure

Two forms of data representation and analysis are reported in this paper. First, descriptive visualisations—these are based

TABLE I
PRODUCT ATTRIBUTES, ASSOCIATED SURVEY QUESTION, AND VARIABLE NOTATION.

Var.	Attribute	Survey Question
a	Visual Appeal	‘How much do you like the look of this product?’
v	Value for Money	‘How happy would you be to pay x for this product?’ (where x is the retail price per item for the product in question).
h	Healthiness	‘How much can this product contribute to a healthy diet?’
t	Taste	‘How much do you like the taste of this product?’
b	Branding	‘How much does the product brand appeal to you?’
e	Ethics	‘How ethical is this product?’
o	Overall Purchase Intention	‘Overall, how likely are you to buy this product?’

on the IAA method [4], which provides a fuzzy set and thus allows a graphic illustration of all response intervals combined as a whole. This gives an idea of overall degree and areas of agreement, along with the degree of uncertainty that is held within the group as a whole. The resulting fuzzy sets also enable subsequent reasoning over the data using fuzzy logic or similarity measures for example—an area not further explored in this paper.

Second, inferential statistics are applied to key characteristics of the interval-valued ratings. In this case, each interval is decomposed into two point-values: the midpoint (m), and the width (w). These independently represent the mean position along, and spread across the response scale, with the latter representing the uncertainty, or range, associated with each response. We then apply linear mixed effects modelling (an extension of linear regression) to assess the influence of each of these characteristics, in relation to each product attribute, upon the corresponding values for overall product purchase intention.²

This approach estimates the contribution of each attribute’s midpoint and width together upon the model outcome variable, i.e. for Model 1—purchase intention rating midpoint, Model 2—purchase intention rating width. These contributions are represented in the form of β weights. The product attribute rating characteristics are entered as fixed effects, alongside two-way interaction terms ($m \cdot w$). These represent combined effects of both response dimensions, which permit the model to identify effects that concern both a given attribute rating’s position and width—e.g. high certainty around disliking a product’s taste may have an opposite effect on overall purchase intention than high certainty around liking a product’s taste.

²Note that higher widths are only possible towards the centre of the scale. That is, as the midpoint approaches the edge of the scale, a wide interval cannot exist. Also, note that while responses were recorded in the range [1, 100], these data were standardised, through conversion to z-scores, before entry into the model.

These models also account for varying baseline purchase intention ratings between participants and between products, through the inclusion of random intercepts.

Two separate analyses are conducted, pertaining respectively to the dependent variables of overall purchase intention position (midpoint), and overall purchase intention uncertainty (width).

Due to the presence of a high number of initial fixed effects (18), an iterative process of backwards stepwise variable reduction was used to ‘prune’ the variables present within each model, leaving only those that were found to contribute significantly to the outcome variable. This was done for the purposes of increasing model interpretability, although it is important to bear in mind that this method leads to inflation of the Type 1 error rate for variables retained in the final model, by comparison with retaining all initial factors—when interpreting results, confidence in the robustness of effects in relation to their reported p -values should therefore be adjusted accordingly.

Refer to Table I for variable notations. The sum of all simple effects for product attribute rating positions (midpoints) is

$$A_m^z = \beta_1^z x_{i,j}^{am} + \beta_2^z x_{i,j}^{vm} + \beta_3^z x_{i,j}^{hm} + \beta_4^z x_{i,j}^{tm} + \beta_5^z x_{i,j}^{bm} + \beta_6^z x_{i,j}^{em} \quad (1)$$

where β is the coefficient, $x_{i,j}^{am}$ is the value m (midpoint) of attribute a (visual appearance) for i (a given participant) and j (a given product), and z reflects the model’s outcome variable, which may be either m (midpoints) or w (widths) of the overall purchase intention.

The sum of all simple effects for product attribute rating uncertainties (widths) is

$$A_w^z = \beta_7^z x_{i,j}^{aw} + \beta_8^z x_{i,j}^{vw} + \beta_9^z x_{i,j}^{hw} + \beta_{10}^z x_{i,j}^{tw} + \beta_{11}^z x_{i,j}^{bw} + \beta_{12}^z x_{i,j}^{ew} \quad (2)$$

where $x_{i,j}^{aw}$ is the width w of attribute a (visual appearance) for i (a given participant) and j (a given product).

The sum of the interactions between the midpoints and widths of the product attributes is

$$A_{mw}^z = \beta_{13}^z (x_{i,j}^{am} \cdot x_{i,j}^{aw}) + \beta_{14}^z (x_{i,j}^{vm} \cdot x_{i,j}^{vw}) + \beta_{15}^z (x_{i,j}^{hm} \cdot x_{i,j}^{hw}) + \beta_{16}^z (x_{i,j}^{tm} \cdot x_{i,j}^{tw}) + \beta_{17}^z (x_{i,j}^{bm} \cdot x_{i,j}^{bw}) + \beta_{18}^z (x_{i,j}^{em} \cdot x_{i,j}^{ew}) \quad (3)$$

Our initial model formula to explain the overall purchase intention midpoints ($\gamma_{i,j}^{Aom}$) and widths ($\gamma_{i,j}^{Aow}$) is then

$$\gamma_{i,j}^{Aoz} = \beta_0^z + A_m^z + A_w^z + A_{mw}^z + \mu_i^z + \mu_j^z + \epsilon_{i,j}^z \quad (4)$$

where z reflects the model’s outcome variable, which may be either m (midpoints) or w (widths), for participant i relating to product j ; β_0 denotes the fixed intercept; μ_i and μ_j denote respective random intercepts for participant and product; and ϵ represents the error. The remaining β terms (within A_m , A_w and A_{mw}) denote the coefficients of the fixed effects of the product attributes.

Each of the initial models, as presented above, was then subjected to a backwards stepwise variable elimination procedure. During this, fixed effects were iteratively assessed and

TABLE II
RESULTS SHOWING SIGNIFICANT EFFECTS OF PRODUCT ATTRIBUTE MIDPOINTS (m), WIDTHS (w) AND TWO-WAY INTERACTIONS ($m \cdot w$) ON MEAN RATING FOR OVERALL PURCHASE INTENTION (om).

Fixed Effects Estimates	β	SE	t	p
Intercept : (0)	.000	.070	.000	> .999
Value m : ($x_{i,j}^{vm}$)	.157	.048	3.262	.001
Healthiness m : ($x_{i,j}^{hm}$)	.092	.042	2.160	.032
Taste m : ($x_{i,j}^{tm}$)	.469	.043	10.982	< .001
Brand m : ($x_{i,j}^{bm}$)	.270	.043	6.343	< .001
Random Effects Estimates μ				
Participant intercept (i)	.303			
Product intercept (j)	.107			
Residual $\epsilon_{i,j}$.622			

N = 320, DF = 315, AIC = 685.1, BIC = 715.1

those that did not significantly contribute to the overall model were removed. Specifically, this process began by selection, from the pool of all non-significant fixed effects, of the effect with the t -statistic closest to zero. This variable was then removed, and the resulting model directly compared with the preceding one, using the Theoretical Likelihood Ratio test. This was implemented through the MATLAB *fitlme* and *compare* functions. If the benefit of retaining the variable in question was calculated to be non-significant, then the model with the lower Bayesian Information Criterion (BIC) was retained into the next iteration. This procedure continued until a final model was determined, within which all fixed effects were statistically significant. Final model parameters were then re-estimated using Restricted Maximum Likelihood (REML) method, to reduce bias in random effect estimates.

III. RESULTS

A. Descriptive Visualisations

Example visualisations of the interval-valued data, modelled as fuzzy sets, using the IAA are shown in Figures 3, 4 and 5.

B. Inferential Statistics

Table II shows the factors retained in the final predictive model for mean position of overall purchase intention rating, subsequent to the stepwise variable removal process. Model results identify four key factors that hold substantial influence over the overall purchase intention rating position. These relate to four product attributes, and all relate to the mean position, rather than the width, of the attribute rating. Of these, taste was found to be the most robust factor, followed by brand, then price, and then perceived healthiness. Neither visual appearance of the actual product nor perceived ethics associated with the product made significant contributions to overall mean purchase intention. Moreover, in this case the width of none of the interval-valued attribute ratings significantly influenced mean purchase intention rating. Furthermore, no significant two-way interaction terms (i.e. $xm \cdot xw$) were found, indicating that the attribute rating widths did not substantially moderate the influence of the attribute rating positions on overall purchase intention.

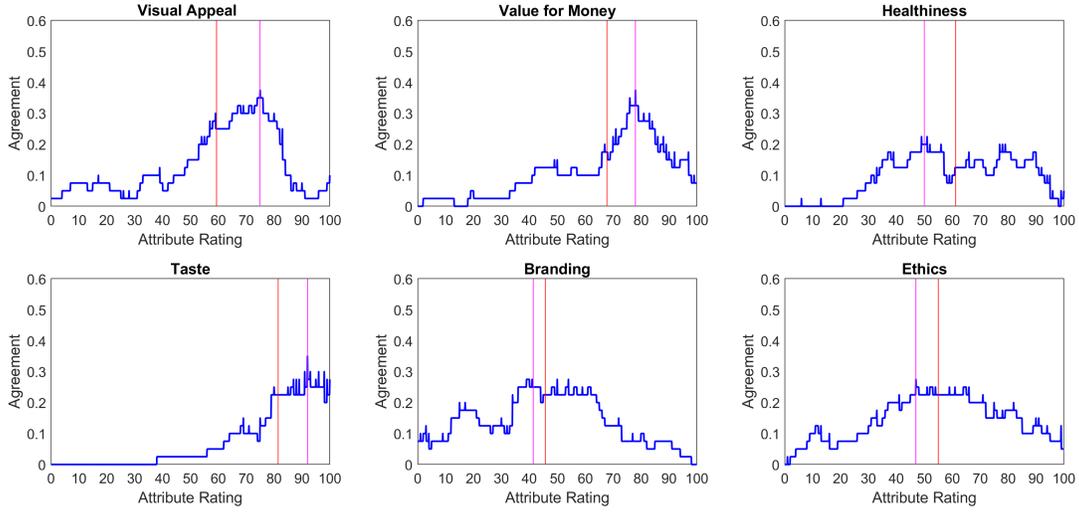


Fig. 3. IAA plots showing subjective importance ratings given for each of the six product attributes, before product exposure. Red vertical lines indicate IAA centroid. Magenta indicate mean of maximum agreement values.

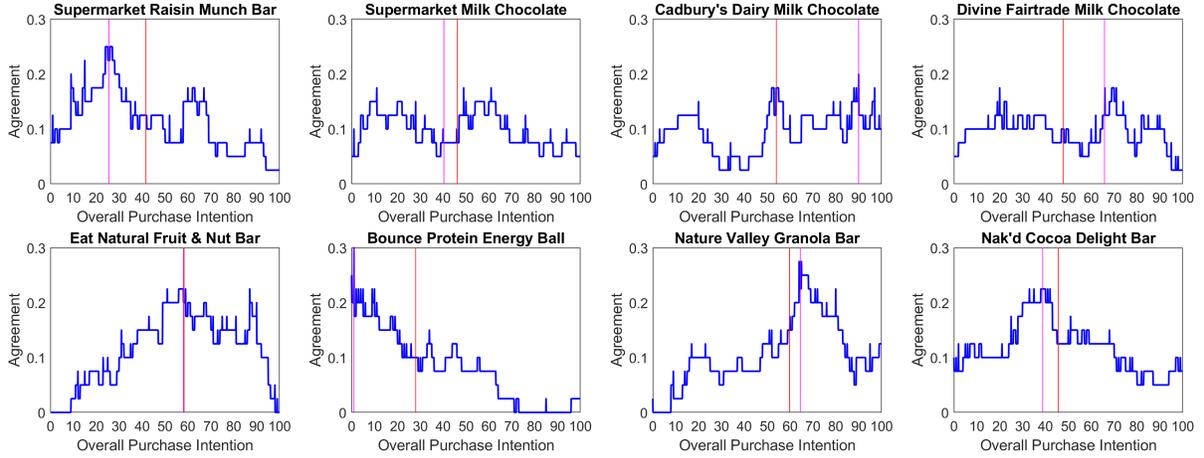


Fig. 4. IAA plots showing overall purchase intention ratings given for each of the eight snack products, after product exposure. Red vertical lines indicate IAA centroid. Magenta indicate mean of maximum agreement values.

Table III shows the factors retained in the final predictive model for uncertainty surrounding overall purchase intention—as represented by rating width—subsequent to the stepwise variable removal process. Model results identify six factors that significantly influenced uncertainty around overall purchase intention. These relate to three different product attributes: visual appearance, taste, and ethics. Three of the six effects relate to mean attribute rating position, one to the width, and two to the interaction of mean position and width (i.e. $xm \cdot xw$). Overall purchase intention was more certain when taste, ethics and visual appearance mean rating positions were low—implying that participants were more certain about their intention to purchase a product when they held negative sentiment rather than positive sentiment toward a product on these three factors. When visual appearance ratings were more certain so were overall ratings. Two-way interaction terms between position and width of ratings for each ethics and appearance show that the combined effects of these differed significantly from the additive effects of

TABLE III
RESULTS SHOWING SIGNIFICANT EFFECTS OF PRODUCT ATTRIBUTE MIDPOINTS (m), WIDTHS (w) AND TWO-WAY INTERACTIONS ($m \cdot w$) ON WIDTHS OF OVERALL PURCHASE INTENTION RATINGS (ow).

Fixed Effects Estimates	β	SE	t	p
Intercept : (\emptyset)	.077	.143	.542	.589
Appearance m : ($x_{i,j}^{am}$)	.114	.037	3.104	.002
Taste m : ($x_{i,j}^{tm}$)	.103	.033	3.116	.002
Ethics m : ($x_{i,j}^{em}$)	.246	.100	2.453	.015
Appearance w : ($x_{i,j}^{aw}$)	.084	.039	2.140	.033
Appearance $m \cdot w$: ($x_{i,j}^{am} \cdot x_{i,j}^{aw}$)	.135	.045	2.963	.003
Ethics $m \cdot w$: ($x_{i,j}^{em} \cdot x_{i,j}^{ew}$)	.280	.110	2.546	.011
Random Effects Estimates				
μ				
Participant intercept (i)	.853			
Product intercept (j)	.077			
Residual $\epsilon_{i,j}$.496			

N = 320, DF = 313, AIC = 634.3, BIC = 671.8

each rating characteristic independently—i.e. effects of one dimension of these factors influenced the other.

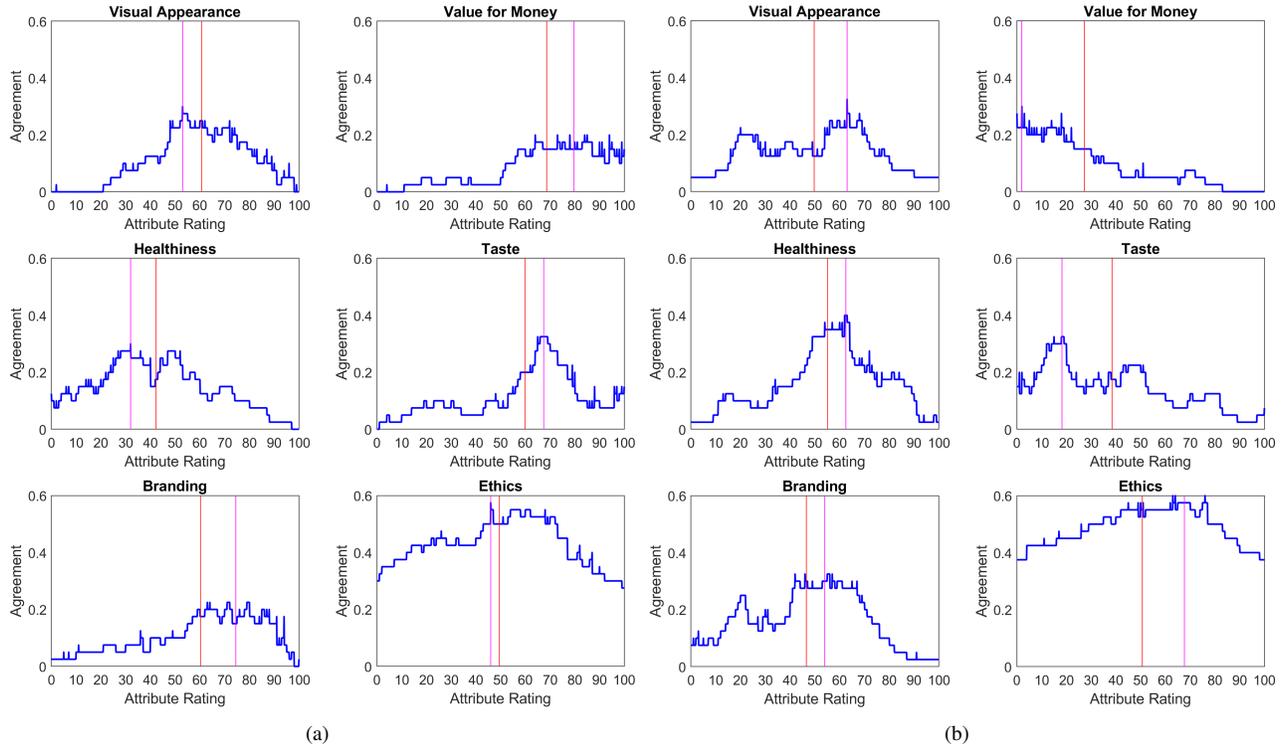


Fig. 5. IAA plots comparing ratings—provided during product exposure—between Cadbury’s Dairy Milk (a), and Bounce Energy Ball (b), for each of the six attributes. Red vertical lines indicate IAA centroid. Magenta indicate mean of maximum agreement values.

TABLE IV

COMPARISON OF AGGREGATED HYPOTHETICAL PRODUCT ATTRIBUTE IMPORTANCE MEASURES—AS PROVIDED IN THE PRE-EXPOSURE SURVEY PHASE—VERSUS MODEL-DERIVED FACTOR EFFECTS.

Importance Measure	<i>a</i>	<i>v</i>	<i>h</i>	<i>t</i>	<i>b</i>	<i>e</i>
Mean of Midpoints	59.6	74.1	61.6	88.1	41.0	59.2
IAA Centroid	59.4	67.8	61.2	81.5	45.7	55.0
IAA Mean of Maxima	75.0	78.0	50.0	92.0	41.4	47.0
LMEM β -value	NS	.16	.09	.47	.27	NS

TABLE V

SUBJECTIVE FEEDBACK RATINGS FOR THE ELLIPSE RESPONSE MODE (RANGING FROM -2.5, 2.5).

Question	Mean	S.E.	95% C.I.
Easy to Use	1.50	.11	.22
Unnecessarily Complex	-1.10	.17	.33
Effectively Communicate	1.25	.15	.29
Overall Liking	1.48	.13	.25

C. Attribute Importance—Subjective Ratings vs Model Effects

We aggregate and summarise hypothetical product attribute importance ratings—as provided by participants in the first stage of the survey (before product exposure), using three different descriptive statistics. We then compare these against the output of the first mixed effects model, which determined the relative influence of attribute ratings, provided during product exposure, on overall purchase intention ratings, provided post-product exposure. The aim was to assess whether there was any substantial discrepancy between subjective and model derived attribute importance. Results are shown in Table IV.

Interestingly, the rank order of participants’ self-reported factor importance ratings, according to both the Mean of Midpoints and the IAA Centroid metrics, was: taste > value > health > appearance > ethics > brand. The only difference for the IAA Mean of Maxima metric was the ordering of appearance and health. All subjective measures agreed that brand was least important factor. By contrast, the order of importance determined by the mixed-effects model was: taste > brand > value > health > appearance ≈ ethics, suggesting

that when predicting actual overall purchase intention ratings, brand rating provided the second greatest contribution.

D. Subjective Feedback

Subjective feedback ratings concerning participant sentiment towards use of the ellipse response mode were obtained, along four key characteristics. These were collected using a conventional five point ordinal scale, ranging from 1—Strongly Disagree, to 5—Strongly Agree. Results, shown in Table V, are re-scaled with zero as the scale midpoint (range -2.5, 2.5), such that negative values show disagreement and positive values agreement. It is clear from the mean and 95% confidence intervals that, as a group, participants rated their agreement as significantly greater than zero on the three positive factors, and significantly lower than zero on the one negative factor.

A linear multiple regression model was conducted to assess the influence of the three former factors on overall liking of the response format, results are shown in Table VI. These indicate that ease of use and how effectively participants perceived the ellipse response mode to allow them to communicate their desired responses were the primary predictors of overall liking.

TABLE VI
EFFECTS OF SUBJECTIVE FEEDBACK CHARACTERISTICS ON THE
DEPENDENT VARIABLE OF OVERALL LIKING OF THE ELLIPSE RESPONSE
FORMAT.

Characteristic	<i>B</i>	<i>p</i>
Intercept	.152	.441
Ease of Use	.590	< .001
Unnecessary Complexity	.031	.708
Effective Communication	.377	< .001

DF = 36, Adjusted R^2 = .618

IV. SUMMARY, CONCLUSIONS AND FUTURE WORK

This paper documents a real-world application of the DECSYS software and interval-valued ellipse response format. We argue that collecting interval-valued responses provides greater information content within each individual response—and that the ellipse response mode permits coherent, efficient and intuitive capture of the uncertainty, variability or vagueness associated with each answer. In this study, we demonstrate the capability of DECSYS to capture interval-valued responses in a relevant practical context—consumer preference research—and evaluate how this may deliver insights into consumer values, product perceptions, and the importance of different factors in forming purchase intentions. This appraisal was designed to provide an initial exposition of the potential value offered by the use of interval-valued data generally, and DECSYS and the ellipse response format specifically, to encourage broader uptake of and engagement with these methods.

In the study, participants sampled a total range of eight snack food products, and rated these on a range of six attributes, which could each potentially have substantial influence on overall sentiment and purchase intention towards the products in question. One primary objective was to determine which factors held substantial influence over overall product purchase intention, across the full range of products—we did not aim to robustly establish superiority or preference for any given product across the broader population. Crucially, as we collected interval-valued ratings, we were also able to assess the influence of uncertainty around product ratings, in terms of both product attributes and overall purchase intention.

Results found that the mean interval-valued ratings representing product purchase intention were substantially influenced by four product attributes: taste, branding, price and perceived healthiness. The uncertainty around these, or other, product attribute ratings was not found to add significant explanatory value to this model. However, the capture of uncertainty associated with overall purchase intentions permitted identification—through another model—of a number of factors that influenced this degree of uncertainty. These included attributes that were not found to significantly affect average purchase intention rating. For example, perceptions around both the appearance and ethics of a product were found to significantly influence uncertainty around overall purchase intention, although they did not affect mean degree of purchase intention.

The capture of interval-valued responses can also provide interesting insights concerning consumer perceptions of different product attributes—as illustrated through fuzzy-set based data visualisations such as the IAA. These illustrate differences between subjective importance ratings provided for the six attributes before product exposure, between overall purchase intention for each of the eight products following product exposure, and between ratings of each of the eight attributes themselves during product exposure. For instance, these clearly show a high relative degree of uncertainty, across a large proportion of respondents, in relation to product ethics. In addition, these plots highlight whether group ratings more closely approximate uni-, bi- or multi-modal distributions. Combined with further analysis—e.g. to break down the characteristics or consumer archetypes of respondents—this may provide further valuable insights into drivers behind consumer sentiment and purchase intention.

Before sampling the products, participants estimated how important each of the six factors were when making a product choice. This allowed comparison between these initial subjective ratings and the importance as estimated by the statistical model, based upon subsequent product and purchase intention ratings. Interestingly, this analysis highlighted one substantial discrepancy between self-reported and model determined factor influence. This was in the importance of branding—which was rated by participants, on average, as the least important of the six attributes when making a product choice, but identified by the statistical model as having the second largest effect on average purchase intention.

As an addendum to the primary study, we also solicited participant feedback regarding their use of the (interval-valued) ellipse response mode, as administered through the DECSYS survey software. We found that user feedback was consistently positive. Participants reported that they found the survey easy to use, that it was not unnecessarily complex, that it allowed them to effectively communicate their desired responses, and that they liked it overall. Of course, these ratings should not be over-interpreted in the absence of comparable ratings for traditional, or other alternative response formats. In the future, we plan to report findings from more comprehensive research designs, which will focus on both user feedback and more objective ease-of-use measures, pertaining to the ellipse response mode and alternatives of varying complexity, in order to empirically inform the ‘effort vs information trade-off’. Nonetheless, the present results are promising in respect to the potential uptake of this response format from the perspective of maintaining a very manageable participant workload.

To conclude, an absence of tools to allow easy collection, collation and analysis of interval-valued responses has held back their use in wider research. The DECSYS software facilitates electronic capture of interval-valued responses, using an efficient and intuitive ellipse response format. This paper documents an initial appraisal of this approach, as applied to the practical research problem of determining factors that influence consumer sentiment and purchase intention—for a range of snack food products. The study demonstrates

insights obtained from interval-valued response data, both using descriptive data visualisations and inferential statistical analyses. Although the statistical methods applied in this study involved both position and width dimensions of interval-valued survey responses, these dimensions were in this case represented independently, by point-values. In the future we plan to develop and apply tailored statistical approaches and tools to enable more holistic collation and analysis of interval-valued responses. Note that due to the sample of the present study—considering both size and location—we do not claim that outcomes are representative of consumer opinion within the general population. Nevertheless, we believe that these findings do provide credible evidence for the importance of further exploring the utility of intervals, and thus reinvesting effort into associated research. We propose that there is substantial scope for such research, pertaining to many aspects of this subject area—from mathematics, statistics, and information theory, to general methodology, and broader qualitative analysis. In terms of fuzzy sets, significant potential exists to leverage both existing tools and future research to support the modelling, representation (as shown here with the IAA) and reasoning with uncertain, interval-valued data.

Future work, some of which is under way, will also be necessary to reinforce and further build upon the foundations of empirical evidence for the value of capturing this type of response-data, both in general and as applied to a variety of specific contexts. Once a critical mass of such evidence is achieved, and coupled with a sufficient degree of exposure and acceptance within the wider research community, it is hoped that the use of interval-valued survey responses will see much broader adoption and provide improvements across a multitude of research domains.

REFERENCES

- [1] Z. Ellerby, J. McCulloch, J. Young, and C. Wagner, “Decsys—discrete and ellipse-based response capture system,” in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–6, IEEE, 2019.
- [2] S. Miller, C. Wagner, J. M. Garibaldi, and S. Appleby, “Constructing general type-2 fuzzy sets from interval-valued data,” in *2012 IEEE International Conference on Fuzzy Systems*, pp. 1–8, IEEE, 2012.
- [3] C. Wagner, S. Miller, and J. M. Garibaldi, “Similarity based applications for data-driven concept and word models based on type-1 and type-2 fuzzy sets,” in *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–9, IEEE, 2013.
- [4] C. Wagner, S. Miller, J. M. Garibaldi, D. T. Anderson, and T. C. Havens, “From interval-valued data to general type-2 fuzzy sets,” *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 2, pp. 248–269, 2014.
- [5] J. McCulloch, Z. Ellerby, and C. Wagner, “On comparing and selecting approaches to model interval-valued data as fuzzy sets,” in *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–6, IEEE, 2019.
- [6] R. Likert, “A technique for the measurement of attitudes,” *Archives of psychology*, 1932.
- [7] R. C. Aitken, “A growing edge of measurement of feelings [abridged] measurement of feelings using visual analogue scales,” 1969.
- [8] Z. Ellerby, J. McCulloch, M. Wilson, and C. Wagner, “Exploring how component factors and their uncertainty affect judgements of risk in cyber-security,” in *International Conference on Critical Information Infrastructures Security*, pp. 31–42, Springer, 2019.
- [9] J. Barreiro-Hurlé, S. Colombo, and E. Cantos-Villar, “Is there a market for functional wines? consumer preferences and willingness to pay for resveratrol-enriched red wine,” *Food Quality and Preference*, vol. 19, no. 4, pp. 360–371, 2008.
- [10] E. Cohen, S. Goodman, L. Casini, and A. M. Corsi, “Consumer preferences of wine in Italy applying best-worst scaling,” *International Journal of Wine Business Research*, 2009.
- [11] M. D’Amico, G. Di Vita, and L. Monaco, “Exploring environmental consciousness and consumer preferences for organic wines without sulfites,” *Journal of Cleaner Production*, vol. 120, pp. 64–71, 2016.
- [12] J. M. Gil and M. Sánchez, “Consumer preferences for wine attributes: a conjoint approach,” *British Food Journal*, 1997.
- [13] U. R. Orth, M. M. Wolf, and T. H. Dodd, “Dimensions of wine region equity and their impact on consumer preferences,” *Journal of Product & Brand Management*, 2005.
- [14] M. Sánchez and J. M. Gil, “Consumer preferences for wine attributes in different retail stores: a conjoint approach,” *International journal of wine marketing*, 1998.
- [15] G. A. Baker, “Consumer preferences for food safety attributes in fresh apples: market segments, consumer characteristics, and marketing opportunities,” *Journal of Agricultural and Resource Economics*, pp. 80–97, 1999.
- [16] S. Denver and J. D. Jensen, “Consumer preferences for organically and locally produced apples,” *Food Quality and Preference*, vol. 31, pp. 129–134, 2014.
- [17] F. R. Harker, F. A. Gunson, and S. R. Jaeger, “The case for fruit quality: an interpretive review of consumer attitudes, and preferences for apples,” *Postharvest biology and technology*, vol. 28, no. 3, pp. 333–347, 2003.
- [18] S. R. Jaeger, Z. Andani, I. N. Wakeling, and H. J. MacFie, “Consumer preferences for fresh and aged apples: a cross-cultural comparison,” *Food quality and preference*, vol. 9, no. 5, pp. 355–366, 1998.
- [19] M. L. Loureiro, J. J. McCluskey, and R. C. Mittelhammer, “Assessing consumer preferences for organic, eco-labeled, and regular apples,” *Journal of agricultural and resource economics*, pp. 404–416, 2001.
- [20] N. K. Novotorova and M. A. Mazzocco, “Consumer preferences and trade-offs for locally grown and genetically modified apples: a conjoint analysis approach,” *International Food and Agribusiness Management Review*, vol. 11, no. 1030-2016-82719, pp. 81–104, 2008.
- [21] A. Hackbarth and R. Madlener, “Consumer preferences for alternative fuel vehicles: A discrete choice analysis,” *Transportation Research Part D: Transport and Environment*, vol. 25, pp. 5–17, 2013.
- [22] J. P. Helveston, Y. Liu, E. M. Feit, E. Fuchs, E. Klampfl, and J. J. Michalek, “Will subsidies drive electric vehicle adoption? measuring consumer preferences in the us and china,” *Transportation Research Part A: Policy and Practice*, vol. 73, pp. 96–112, 2015.
- [23] C. D. Higgins, M. Mohamed, and M. R. Ferguson, “Size matters: How vehicle body type affects consumer preferences for electric vehicles,” *Transportation Research Part A: Policy and Practice*, vol. 100, pp. 182–201, 2017.
- [24] W. Ko and T.-K. Hahn, “Analysis of consumer preferences for electric vehicles,” *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 437–442, 2013.
- [25] F. Liao, E. Molin, and B. van Wee, “Consumer preferences for electric vehicles: a literature review,” *Transport Reviews*, vol. 37, no. 3, pp. 252–275, 2017.
- [26] C. Morton, J. Anable, and J. D. Nelson, “Exploring consumer preferences towards electric vehicles: The influence of consumer innovativeness,” *Research in transportation business & management*, vol. 18, pp. 18–28, 2016.
- [27] E. Menichelli, H. Kraggerud, N. V. Olsen, and T. Næs, “Analysing relations between specific and total liking scores,” *Food quality and preference*, vol. 28, no. 2, pp. 429–440, 2013.
- [28] D. D. Rucker, Z. L. Tormala, R. E. Petty, and P. Briñol, “Consumer conviction and commitment: An appraisal-based framework for attitude certainty,” *Journal of Consumer Psychology*, vol. 24, no. 1, pp. 119–136, 2014.
- [29] E. Maier, R. Wilken, and F. Dost, “The double benefits of consumer certainty: combining risk and range effects,” *Marketing Letters*, vol. 26, no. 4, pp. 473–488, 2015.
- [30] R. E. Petty, P. Briñol, and Z. L. Tormala, “Thought confidence as a determinant of persuasion: the self-validation hypothesis,” *Journal of personality and social psychology*, vol. 82, no. 5, p. 722, 2002.
- [31] S. R. Gross, R. Holtz, and N. Miller, “Attitude certainty,” *Attitude strength: Antecedents and consequences*, vol. 4, pp. 215–245, 1995.
- [32] J. Brooke et al., “Sus-a quick and dirty usability scale,” *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.