ABSTRACT

Choice experiments (CE), involving multi-attribute choices, are increasingly used in economics to value non-marketed goods. Such choices require individuals to process large amounts of information, shown to trigger partial information strategies in participants. We develop a new framework in which information processing is treated as a latent (unobservable) process. Testing our approach by combining CE and visual attention (VA) data gathered from eye-tracking, we show that treating information processing as a latent process (LIP) outperforms models assuming full information processing (FIP) or binary information processing (BIP). Our modelling of VA results in a number of key findings. We show that the relationship between VA and individuals' preferences depends on the type of product attribute. More specifically, preferences for "easier to process" attributes appear to be less influenced by changes in underlying level of VA than "harder to process" attributes. In turn this impacts on willingness-to-pay estimates, with the LIP model resulting in smaller values than those obtained with the FIP model. Our results have implications for CE designers. More time should be spent getting subjects to understand more complicated attributes of the CE. Our results are likely to extend beyond experimental choices (stated preferences) to actual choices (revealed preferences).

Keywords: Choice experiment; Stated preferences; Eye-tracking; Information processing; Choice modelling

APA Codes: 2340, 2346, 3920
JEL codes: I12, D80, C35

1. Introduction

Choice experiments (CEs) are commonly used in economics to value products or services when markets for them do not exist (de Bekker-Grob, Ryan, and Gerard, 2012; Hoyos, 2010; Lagarde & Blaauw, 2009). Developing from the disciplines of psychology (Luce, 2005; Thurstone, 1927) and economics (Lancaster, 1966; Manski, 1977), CEs present participants with the task of choosing one option from two or more hypothetical products, each described in terms of a set of attributes. Based on the Lancasterian theory of demand, the CE stipulates that the utility of a product comes from its features rather than the product itself, what is known as multi-attributes utility (Lancaster, 1966). Whilst this approach measures the relative contribution of different product features to individuals' choices, as well as trade-offs between attributes, it raises important questions regarding how individuals process large amounts of information. Previous studies have investigated information processing (IP) in the context of multi-attributes choices by determining which product attributes do not contribute to individuals' choices, what is known as "attribute non-attendance" (ANA) (Hensher & Greene, 2010; Carlsson, Kataria, and Lampi, 2010; Hole, 2011; Hole, Kolstad, and Gyrd-Hansen, 2013). These studies indicate that ANA has important consequences on modelling of discrete choices, computation of willingness-to-pay for product improvement and predictions of market shares. However, due to lack of data about the cognitive processes taking place during individuals' choices, these studies provide no explanation of how ANA occurs. For example, does ANA occur early in the decision making process when individuals start processing or collecting information about the products, or at a later stage when individuals choose their *preferred* product?

We explore the role of visual attention (VA) in processing multi-attribute information, and how this relates to individuals' choices. We use an eye-tracker to explore how individuals visually process the information. An eye-tracker records participants' eyes movements, thus providing data on VA, such as fixation times, while making multi-attributes choices. Using an eye-tracker within a CE offers a unique setting to simultaneously observe both processes and outcomes of individuals' choice behaviour. To the best of our knowledge, this experimental setting has only been used in Balcombe et al (2015) who analysed the impact of partial information processing (referred to as "visual ANA") on preferences for food labelling. They used a "shrinkage approach" to represent the relationship between visual attention and ANA, allowing for a "*continuum between two extremes from where [visual] non-attendance is non-informative to the case where non-attendance signifies that the non-attender has zero marginal utility for the attribute in question*". This approach showed a weak relationship between changes in VA and marginal utilities. We argue that their result may be due to the procedure they used to identify visual ANA (see section 3.2). Key to their approach is the use of a cut-off point to convert continuous measures of VA, such as fixation times (in milliseconds) on different attributes, into a binary measure of visual ANA. Whilst this binary approach is conceptually *straightforward* to implement and compatible with previous literature on ANA; it is likely to provide an incomplete account of the role of VA in multi-attribute choices.

Commented [nk1]: Reply #2

Our study contributes to the literature by exploring alternative ways of accounting for VA in the measurement of preferences. Whilst ANA would be associated with one extreme (binary) form of multi-attribute information processing (i.e., *is the piece of information ignored vs. considered?*), our approach is based on the assumptions that (i) attention would be best described as a continuous rather than a binary concept, and (ii) recognising the continuous nature of attention would reveal more information about individuals' preferences. Following Shimojo et al (2003) and Krajbich & Rangel (2011), VA and preferences cannot be perfectly dissociated as they influence each other in a positive feedback loop i.e., "*The longer I look at something, the more I like it*" and "*The more I like something, the longer I look at it*". This suggests that VA contains useful information about individuals' preferences. We develop a framework in which measures of VA are considered as indicators of an underlying level of "information intake/processing" (i.e., *how much individuals learn about the multi-attribute content of the product*). Our results indicate that measures of VA improve the modelling of choice behaviour. Our study also contributes to the literature by showing that the relationship between VA and individuals' preferences depends on the type of product attribute. More specifically, preferences for "easier to process" attributes appear to be less influenced by changes in underlying level of VA than "harder to process" attributes. In turn this impacts on willingness-to-pay estimates. Our results have important implications for the analysis of choice data (revealed and stated) as well as the design of CEs.

Commented [nk2]: Reply #1

The rest of the article is organised as following. In section 2, we describe the CE. In the section 3, we develop the analytical framework for investigating the effect of VA on individuals' multi-attribute choices. In sections 4 we present and discuss the results. Section 5 makes concluding comments.

2. Eye-Tracking Choice experiment

2.1 Sample of Participants

58 psychology students were recruited using online advertisement (See Table 1 for a description of sample characteristics). Recruitment was on a first-come-first-served basis and took part in return for course credit. The study protocol was approved by ethics committee of the School of Psychology at the University of Aberdeen (UK).

2.2 Choice Experiment

Participants' choices were recorded for an existing CE concerned with preferences for characteristics of health and lifestyle (H&L) programmes (Ryan et al, 2015). Each H&L programme was described by seven attributes: type of H&L programme [PROG]; objective of the programme [GOAL]; level of weight reduction [WEIGHT]; level of high blood pressure risk reduction [HBP]; level of diabetes (Type 2) risk reduction [DIABETES]; time commitment to the programme [TIME]; and cost of the programme [COST] (Table 2). Twelve choice tasks were derived using experimental designing procedures, allowing main effects of attributes to be estimated. Choices were presented on a computer; participants were asked to select their most preferred alternative in each task (Figure 1). For each trial, participants looked at the choices until they reached a decision, and then pressed the space-bar on the keyboard to reveal the mouse cursor, indicating that they were ready to indicate their response. With the mouse, they then clicked on their preferred H&L programme. If they did not prefer any of the programs, participants had the option to click outside the image to indicate they preferred to stick with their current exercise and diet program. Before the 12 experimental tasks, four practice tasks were presented to familiarise participants with the general layout of the screen and task. These tasks were excluded from analysis[1]. The order of both the choice tasks and alternatives within the choice tasks were randomised across participants to minimise potential ordering effects. No time limit was imposed on responses.

The participants were allocated to one of two experimental conditions. In the first (N_1=28), attributes were presented in the following order (from top to bottom): PROG, WEIGHT, GOAL, DIABETES, HBP, TIME, and COST. In the second (N_2=30), the order of the attributes was reversed (i.e., COST; TIME; …; WEIGHT; PROG). Between the two groups, we also swapped the two choice alternatives (left and right). The two experimental conditions were otherwise identical. This experimental manipulation aimed to minimise the influence of presentation effect on respondents' choices.

2.3 Eye-Tracking

While participants completed the CE on a computer in a dedicated eye-tracking laboratory, their eyes' movements were recorded using an eye-tracker (i.e., EyeLink 1000 system) calibrated for each subject individually with the eye tracker's default 9-point calibration procedure. Eyes' movements were recorded at a 1,000 Hertz rate (i.e., 1 observation every 1 millisecond). Participants were seated at a distance of approximately 77 cm from the eye-tracker. To avoid large head movements that may interfere with accurate eye tracking a combined head-and-chin rest was used. Each trial started with a fixation point presented in the middle of the screen. Participants were asked to fixate this stimulus, after which the experimenter initiated the trial.

Eyes' movements were automatically divided into fixations and saccades using the default algorithm of the eye tracking system. While the eye remains relatively still during fixations, it moves at high speeds during saccades (Duchowski, 2007). It may be reasonably assumed that, for our setup, information processing only took place during fixations, and therefore only these observations were used to measure VA. We further assumed that

[1] The 12 experimental trials were intermixed with two non-experimental trials used to test Sen's expansion and contraction properties of individuals' choices. As these two additional tasks included three choice options, instead of two, they were also excluded from the analyses.

information processing could not take place with fixations shorter than 50 milliseconds. The remaining fixations were analysed in terms of where they were directed to on the display. To this end, 24 regions of interest (ROI) were drawn around the 24 boxes in each choice display, and fixations were assigned to the box they were directed to (Figure 1).

After excluding fixations from the four practice tasks and those outside of the 24 ROIs, 37,784 fixations remained. We further excluded 5,822 fixations (15.4%) on the descriptive column (ROI 1 to 8) or alternative labelling (ROI 9 and 17) because they provide no meaningful information about the multi-attributes content of the H&L programmes. We also excluded 728 fixations (2.3%) with a fixation time shorter than 50 milliseconds. This resulted in 31,234 valid fixations (observations) for analysis.

3. Analysis

Multi-attributes choices are typically analysed within the random utility maximisation (RUM) framework. This framework can be seen as a collection of assumptions about participants' choice behaviour: utility maximisation; random utility (Thurstone, 1927); and attribute-wise utility (Lancaster, 1966). Following the utility maximisation hypothesis, participants' choices are the outcomes of a process comparing the latent utility (U) of the different alternatives on offer. The random utility hypothesis stipulates that the utility of each alternative can be divided into a systematic/observable utility (V) and a random/unobservable utility (ε). Because of this random component, the participants' choice behaviour becomes probabilistic and then we analyse the probabilities of making particular choices (P). The attribute-wise utility hypothesis is used to define the observable portion of the utility as a combination of preferences (β) and attributes' values (X). Usually the indirect utility function (V) is assumed to be linear and additive:

$$U_{njt} = V_{njt} + \varepsilon_{njt} \qquad (1)$$
$$V_{njt} = \sum_k \left(\beta_k X_{knjt}\right) \qquad (2)$$

Where participants are indexed by n=(1,…,N), alternatives by j=(1,…,i,…,J), (i) denotes the selected alternative, choice tasks by t=(1,…,T), and attributes by k=(1,…,K).

3.1 The Full Information Processing (FIP) model

Based on the RUM framework, different choices models can be derived by making further assumptions about the behaviour of the error terms. In practice (ε) are often assumed to be independently and identically distributed (IID) as type I extreme values, leading thus to the multinomial logit (MNL) model (McFadden, 1974; Train, 2009).

$$P_{nit} = \frac{\exp(\mu_{nit} V_{nit})}{\sum_j \exp(\mu_{njt} V_{njt})} \qquad (3)$$

The (μ) parameter indicates the scale of utility (or relative size of the unobservable utility component), which is inversely related to the variance of model errors (σ_{ε}^2) and is usually normalised to 1 for identification purpose. This scale parameter can also be specified as a function of observable characteristics (Z) that would capture systematic changes in choices consistency, leading thus to a scaled (or heteroscedastic) MNL model (SMNL)[2].

$$\mu_{njt} = \exp(\gamma' Z_{njt}) \qquad (4)$$

Several studies have followed this approach to explore the effects of choice task properties, such as location in the sequence of tasks (e.g., 1st vs last task answered by the participants), on choices consistency (Caussade et al, 2005; DeShazo & Fermo, 2002; Scarpa et al, 2011), or the effects of respondents' characteristics, such as experience and knowledge of the goods (Czajkowski, Hanley, and LaRiviere, 2016; Czajkowski, Hanley, and LaRiviere, 2015; LaRiviere et al, 2014). In our study, one could expect choices consistency to differ between the two experimental conditions and then we specify the scale parameter as a function of attributes ordering (ORDER). Kjaer et al (2006) investigated whether displaying the COST attribute either at the top or bottom of the choice options would have a significant impact on its marginal utility, and they showed higher errors variance when the COST was displayed at the top.

Commented [nk3]: Reply #5

This initial formulation of the RUM model makes two restrictive assumptions about choice behaviour. First, all participants are assumed to share the same preferences for attributes (i.e., $\beta_{nk} = \beta_k \forall(n)$). This assumption of preferences homogeneity neglects inter-individuals differences in tastes. Second, as implied by the (Σ) operator in equation 2, participants are assumed to consider all the attributes when making their choices. As indicated by the ANA literature, this assumption of full information processing (FIP) may be unrealistic because participants' decisional resources are limited and they may then adopt simplifying decision rules to decrease the cognitive burden of the choice tasks (e.g., to consider only the price and quality attributes).

We estimate this homogeneous FIP model as a 1st benchmark model (FIP-I) to assess the benefits of alternative (more realistic) choice models. The indirect utility function (V) was specified as:

[2] In the choice modelling literature the concept of choice consistency has been assimilated to the variance of the unobservable component (σ_{ε}^2) in the utility function. Intuitively when participants are not consistent in their choices (i.e., high σ_{ε}^2), for example by making random choices, the size of the unobservable component (ε) becomes large relative to the observable component (V), and then the attributes influence on individuals' choices tend to be null.

$V_{ntj} = \exp(\gamma ORDER_{ntj})(\beta_0 NO_{ntj} + \beta_1 PROG_{ntj} + \beta_2 WEIGHT_{ntj} + \beta_3 GOAL_{ntj} + \beta_4 DIABETES_{ntj} + \beta_5 HBP_{ntj} + \beta_6 TIME_{ntj} + \beta_7 COST_{ntj})$ (5)

More information about the model variables (attributes) can be found in Table 2. The ORDER variable refers to the order in which attributes were presented in the choice options (i.e., initial vs. reversed order). The NO variable refers to choosing not to have an H&L programme and the corresponding parameter (β_0) captures respondents' preference for this opt-out option.

Commented [nk4]: Reply #4

We then relax the assumption of preferences homogeneity by specifying (β's) as random parameters (McFadden & Train, 2000). In this heterogeneous version of the FIP model, participants' preferences are described by a distribution of preferences over the sample. In line with literature, we assume that participants' preferences follow a multi-variate normal distribution with mean vector (μ_K) and diagonal covariance matrix (σ_K).

Commented [nk5]: Reply #3.4

$V_{ntj} = \exp(\gamma ORDER_{ntj})[\beta_0 NO_{ntj} + (\mu_1 + \sigma_{1n})PROG_{ntj} + (\mu_2 + \sigma_{2n})WEIGHT_{ntj} + (\mu_3 + \sigma_{3n})GOAL_{ntj} + (\mu_4 + \sigma_{4n})DIAB_{ntj} + (\mu_5 + \sigma_{5n})HBP_{ntj} + (\mu_6 + \sigma_{6n})TIME_{ntj} + (\mu_7 + \sigma_{7n})COST_{ntj}]$ (6)

The model (simulated) log-likelihood is:

$SLL(\mu, \sigma, \tau) = \sum_n \ln\left(\frac{1}{R}\sum_r \prod_t P_{nti}(\beta_n^r)\right)$ (7)

Where r=(1,…,R) denotes Halton draws of (β_n^r) from the multi-variate normal distribution[3].

We estimate the heterogeneous FIP model as a 2nd benchmark model (FIP-II). Given the random parameters model has been described as the most flexible discrete choice model, it is expected to provide an upper bound on how well we can measure variability in participants' choices. At the opposite, the homogeneous formulation of the FIP model appears to be the least flexible approach and then should be the lower bound. Having these two boundaries will help determine to what extent variability in participants' choices can be explained by changes in information processing. Differences in VA are likely to be only one source of the total between-subjects variability in choices. Other factors, such as true heterogeneity in preferences or differences in ability to choose, are also expected to impact respondents' choice behaviour.

[3] Due to the simulation procedure, the model log-likelihood is no longer globally concave and model estimates become sensitive to specification of starting values and number of draws. We checked the robustness of our results by testing 100 different sets of starting values (obtained by perturbations of initial FIP estimates) and by estimating the random parameters model with increasing number of Halton draws (from 300 to 3,000). Detailed results can be obtained from the corresponding author on request.

3.2 The Binary Information Processing (BIP) model

Collection of eye-tracking data alongside our CE enables direct observation of whether (and how) participants visually process multi-attribute information. For instance an attribute that has not been looked at is unlikely to influence participants' choices. Whilst attribute non-attendance (ANA) is a binary event (i.e., the attribute is either considered (=1) or ignored (=0)), VA is a continuous event (e.g., length of fixation time (FT) on the attribute). The question is then how to convert a continuous measure of information processing into a restrictive binary measure of attendance. This is an important issue as different specifications of the information-attendance link imply different assumptions about the nature of individuals' information processing strategies and preferences.

The simplest way of converting VA into ANA measures is to assume that a particular attribute (k) of an option (j) is considered by the participant (n) in task (t) when its level of VA is above an arbitrary specified threshold (α_k).

Commented [nk6]: Reply #3.5

$$\mathrm{ANA}_{\mathrm{ntjk}} = \begin{cases} 1, \text{if } \mathrm{VA}_{\mathrm{ntjk}} \geq \alpha_k \\ 0, \text{otherwise} \end{cases} \qquad (8)$$

In Balcombe et al (2015) visual ANA was operationalised following a 2-step strategy: Step #1 - in a given choice task (t) and option (j), an attribute (X_{ntj}) was deemed to be visually non-attended by respondent (n) when being fixated less than twice (i.e., $\alpha_k = 2\ \forall k$); Step #2 - the respondent was further classified as a visual non-attender when the attribute (X_{ntj}) was visually attended in less than half of the choice tasks (T=12). We use a different decision rule to convert VA into binary measures of ANA: $\alpha_k \geq 200$ ms $\forall(k)$. We also check the robustness of the BIP model results to changes in the specification of the cut-off points. We assumed different cut-off values from 0 to 2,000 ms by step of 50 ms.

Subsequently the indirect utility function of the BIP model becomes:

$$\mathrm{V}_{\mathrm{ntj}} = \exp(\gamma \mathrm{ORDER}_{\mathrm{ntj}}) \left(\beta_0 \mathrm{NO}_{\mathrm{ntj}} + \textstyle\sum_k \beta_k \left(\left(1 - \mathrm{ANA}_{\mathrm{ntjk}}\right) \mathrm{X}_{\mathrm{ntjk}}\right)\right) \qquad (9)$$

The BIP approach possesses a number of limitations. First, it assumes an error-free (or deterministic) relationship between VA and attribute attendance. When visual ANA measures are allowed to directly modify the matrix of attributes' levels (X), this makes the implicit assumption of a perfect relationship between information processing (i.e., visual search behaviour) and individuals' preferences (i.e., choice behaviour) by constraining the marginal utility of the attribute to be null. That is: *Not looking at an attribute (ANA=1) necessary implies not valuing it (β=0)*. However this is not necessarily the case as one can miss an attribute, because of low visual saliency, but still value it. This might also happen when individuals use already processed information to make inferences about unseen attributes (e.g., "*A product of high quality must be expensive*", such that the individuals will consider the product price for their decision making without having to look at it). In Balcombe et al (2015) this assumption of deterministic relationship was relaxed by using a "shrinkage approach" in which the effect of visual ANA on individuals' preferences was governed by a shrinkage parameter (τ) bounded between 0 and 1. For (τ=0), visual ANA would provide all the information about individuals' preferences i.e., Not looking at an attribute implies null marginal utility; at the opposite end (τ=1), visual ANA would convey no additional information about individuals' preferences.

Second, depending on the choice of the cut-off point, the BIP approach is likely to underestimate either the true proportion of attributes not considered (i.e., low specificity) or the true proportion of attributes considered (i.e., low sensitivity). For example, there is no guarantee that attributes with higher levels of VA (say 250 ms) will be considered when making choices; hence the BIP approach would have low specificity. Assuming that individuals are more likely to consider an attribute when they spend more time looking at it, it would be possible to overcome this limitation by increasing the threshold (e.g., α = 1,000 ms). However some attributes would then be wrongly identified as not considered, leading thus to a low level of sensitivity (i.e., failure to define an attribute as being considered when it is actually the case). Further, given the eye-tracker provides detailed measures of VA (e.g., length of fixation time in milliseconds), the dichotomisation of VA (i.e., missed vs attended) represents a substantial loss of information about information processing.

Third, the BIP approach does not adequately explain how (and why) individuals' VA evolves over time. The use of threshold(s) implies breakings in individuals' choice behaviour and cannot explain why individuals would spend more time than "needed" to process information about multi-attributes content (*If 200 ms are enough to identify the value of an attribute, why do we observe longer fixation times?*). In addition, finding the right threshold value is difficult (if not impossible) as it is likely to be both attribute- and individual-specific, as individuals are known to have different information processing abilities (Nunez, 2015; Ambinder & Simons, 2010; Valuch et al, 2015). For example, a short fixation time (say α = 200 ms) may be sufficient to process information about the cost of the H&L programme, but higher threshold value (say α = 500 ms) might be required for text-based attributes such as type of H&L programme. In addition to varying across attributes and individuals, the cut-off points could also vary across choice tasks because individuals learn about the multi-attribute content and presumably become better at process the information. However it would be practically impossible to identify such highly flexible cut-off points.

Finally, the BIP approach is based on a one-to-one relationship between VA and attendance for the same attribute, and is not influenced by VA on other attributes (e.g., COST attendance only depends on whether the COST attribute was looked at or not). However, VA would be best conceived as a limited resource and then the different attributes are in competition for it, such that attribute attendance should be influenced by attention devoted to all the attributes (Reutskaja et al, 2011; Yang, Toubia, and de Jong, 2013; Sims, 2006; Gabaix et al, 2006).

Given these limitations, we estimate the BIP model as a simplistic approach to describe the effect of multi-attributes information processing on participants' choices. Comparisons with the BIP model will allow determining whether our alternative latent model provides a better account of the attention-attendance link.

3.3 The Latent Information Processing (LIP) model

We propose an alternative approach to represent the role of information processing in multi-attribute choice behaviour. Our approach is based on the central assumption that information processing can be best described as a latent quantity with measures of VA being the indicator. We refer to this as latent information processing (LIP). Individuals' eye movements are expected to reflect general attitudes regarding how much individuals want to learn/process information about the H&L services.

This approach is in line with cognitive psychology and neuro-economics where measures of decision processes (e.g., eye movements) are treated as indicators of cognitive processes (Cavanagh et al, 2014; Krajbich & Rangel, 2011; Liechty, Pieters, and Wedel, 2003). For example, Liechty et al (2003) used a hidden Markov model to analyse eye movements in an advertising viewing task. They hypothesised that eye movements could be used to identify two attentional states, known as global and local attention. These two attentional states, which cannot be directly observed, imply different objectives when processing information (i.e., scanning information vs acquiring information). Liechty et al (2003) used systematic changes in eye movements, such as saccades length, as indicators of the probability of being in a given latent state.

We use a hybrid choice (HC) structure to implement our LIP approach (Ben-Akiva et al, 2002). The LIP model takes the form of a structural equation model (see Figure 2, and equations 11:13)[4]. The "information processing" component (LHS, Figure 2) indicates that the LIP state varies across individuals as a result of differences in personal characteristics (Z). For example, differences in their decision making objectives (e.g., random decision, utility maximisation, regret minimisation, "satisficing" decision, etc.). Whilst the LIP state cannot be directly observed, it can be inferred from measures of VA. Assuming that VA is related to information processing, the LIP state can be interpreted as level of information uptake (i.e., "*how much individuals learn about the programmes*"). When participants look longer at the multi-attributes content, they are expected to become more knowledgeable about the value (utility)

[4] As noted by one of the reviewers, this approach has been also used by Hess & Hensher (2013) to investigate the effect of stated ANA (i.e., at the end of the CE participants were asked to state which attributes they took into consideration when making their choices) on choice consistency (as approximated by changes in errors variance).

of the choice options. Therefore a higher LIP score will be associated with an increase in VA (i.e., longer fixation times on the different attributes) (**H1:** $\boldsymbol{\gamma}' \geq \mathbf{0}$).

The "multi-attribute choice" component (RHS, Figure 2), defines the utility of the choices. Information processing and choice behaviour are linked by allowing the LIP state to influence individuals' preferences for the attributes. Given that LIP reflects information uptake, we anticipate a significant effect of LIP on the strength of preferences for the different attributes (**H2:** $\boldsymbol{\alpha}' \neq \mathbf{0}$).

We further hypothesise that the LIP effect will differ across the attributes, depending on their nature/format. For example, risk information is known to be difficult to process (Fagerlin, Zikmund-Fisher, and Ubel, 2011; Osimani, 2012; Keller et al, 2014). The influence of risk attributes on individuals' choices is thus expected to be more sensitive to changes in level of information intake: the more attentive individuals are, the more likely they become to overcome the information processing difficulties and then to consider the risk attributes in their decisions. Following the processing fluency hypothesis, the ease of information processing could be interpreted by the individuals as a signal about the importance/quality of the attributes (Brakus, Schmitt, and Zhang, 2014; Shah & Oppenheimer, 2007). If an attribute can be easily processed, then individuals would give more weight to that particular attribute in their choices. In our study, ease of information processing can be determined along two dimensions (see Figure 3): (1) "Precision" (i.e., how well defined the attribute is), and (2) "Format" (i.e., whether the information is quantitative [numerical] or qualitative [textual]).

Less precise attributes are prone to interpretation effects and extra effort is required to determine their meaning (Keller & McGill, 1994). In our study, the GOAL attribute describes the objective of the H&L programme, which can be either "Feeling better", "Looking better" or both of them. However "Feeling better" is likely to mean different things to different individuals and then participants have to take an extra step to make the information meaningful to them. *Ceteris paribus* less precise attributes (i.e., HBP; DIABETES; PROG; GOAL) are expected to be more sensitive to changes in level of information intake (LIP) than more precise attributes (i.e., TIME; COST; WEIGHT).

Regarding the attribute format, we expect quantitative (numeric) attributes (i.e., TIME; COST) to be more sensitive to changes in LIP than more qualitative (text) attributes (i.e., PROG; GOAL). Difficulty to process numerical information is usually associated with numeracy (i.e., the individuals' ability to reason and to apply numerical concepts). Previous studies showed that on average individuals are not confident with numerical information and then their decision making was biased towards non-numerical information (Keller, Siegrist, and Visschers, 2009; Gaglio, Glasgow, and Bull, 2012; Keller et al, 2014; Peters et al, 2014; Jasper, Bhattacharya, and Corser, 2016). Therefore a change in level of visual attention devoted to text-based attributes is expected to have a weak-to-moderate effect on individuals' choices.

Also we expect differences in attribute format to have a larger impact than differences in attribute precision.

In summary, the two risk-related attributes, DIABETES and HBP, should be more sensitive to changes in the level of information uptake, followed by the TIME and COST attributes, and by the three remaining attributes (i.e., WEIGHT; PROG; GOAL). The strength of the effect of information uptake on individuals' preferences can be measured with the (α_k/β_k) ratio (r_k) which indicates how large is the marginal effect of information uptake (α_k) compared to marginal utility (β_k) for the attribute (k), and then can be used to make comparisons across the attributes (**H3:** $\mathbf{r_{HBP} = r_{DIABETES} > r_{TIME} = r_{COST} > r_{WEIGHT} \geq r_{PROG} = r_{GOAL}}$)[5].

The LIP model is described in the following equations:

$$IP_n = f_1(Z_n; \lambda) + \eta_n \quad (10a)$$

$$\eta_n \sim N(0,1) \quad (10b)$$

$$VA_{nk} = f_2(FT_{nk}, IP_n; \gamma) + \omega_{nk} \quad (11a)$$

$$\omega_{nk} \sim N(0, \sigma_k) \quad (11b)$$

$$U_{ntj} = f_3(X_{ntj}, IP_n, Z_{ntj}; \beta, \alpha, \tau) + \varepsilon_{ntj} \quad (12a)$$

$$\varepsilon_{ntj} \sim \text{IID EV1} \quad (12b)$$

where parameters: (λ) capture the effect of individuals' characteristics on LIP; (γ) the association between LIP and measures of visual attention; (β) individuals' preferences for attributes; (α) the marginal change in individuals' preferences due to LIP; and (τ) the effect of attributes ordering (i.e., initial vs reversed) on choices' consistency.

We further specify the indirect utility function $f_3(.)$ as linear and additive such that:

$$U_{ntj} = \exp(\tau ORDER_{ntj})\left[\textstyle\sum_k(\beta_k + \alpha_k IP_n)\, X_{ntjk}\right] + \varepsilon_{ntj} \quad (13)$$

This function is comparable with FIP-I model, except that now the latent IP is allowed to influence participants' preferences for attributes through (α_k). As in Balcombe et al (2015), this approach relaxes the assumption of deterministic relationship between VA and choices by allowing non-null VA to convey no additional information about individuals' preferences ($\alpha_k = 0\ \forall(k)$), and in this case the LIP model collapses to the homogeneous FIP model.

The HC model (simulated) log-likelihood can be written as following:

$$SLL(\lambda, \gamma, \sigma, \beta, \alpha, \tau) = \sum_n \ln\left(\frac{1}{R}\sum_r \prod_k \Phi\big(VA_{nk}(\eta_n^r)\big) \prod_t P_{nti}(\eta_n^r)\right) \quad (14)$$

[5] We used ratios instead of marginal effects (α) because the latter provide misleading results as the size of the effect also depends on the size of the marginal utility for the attribute, defined by the measurement scale. For example, the marginal utility for COST is expected to be smaller than the marginal utility for GOAL because their measurement units are different (£1 vs "Feeling good").

where (Φ) denotes the normal probability density. For model estimation, we used no personal characteristics (i.e., Z_n=0) and VA was approximated by the mean centred version of fixation times measured at the individual level[6].

Our proposed LIP approach improves on the BIP model in a number of ways. It does not assume a deterministic relationship between VA and individuals' preferences, allowing for an attribute not to be visually processed without forcing the corresponding preference to be null (and reciprocally an individual may hold null preference for an attribute but still look at it). It also preserves the continuous nature of VA, thus minimising the information loss which occurs with cut-off point(s) and preventing the need to define cut-off points. As detailed in Holmqvist (2011), different measures of VA provide different insights into individuals' decision making processes. Although fixation counts (i.e., the number of times an attribute is looked at), as used in Balcombe et al (2015), and fixation times are highly correlated, the two measures are conceptually different. Fixation times seem to be more reliable for exploring partial information processing in multi-attribute choices because (1) they are less subject to measurement errors and more homogeneous than fixation counts[7], and (2) they allow going beyond *simple* attributes non-attendance (ANA) by considering not only whether an attribute is visually processed or not but also how long it is looked at. Finally, in the LIP model individuals' preferences are not directly impacted by the amount of VA devoted to a particular attribute but by the underlying level of information intake, which is approximated by fixation times. To some extent, this allows, for example, VA_{COST} to influence preferences for TIME (in addition to VA_{TIME}). Our LIP approach is also more robust to endogeneity in measurement of VA. Given VA and individuals' preferences for the attributes cannot be seen as independent, errors in individuals' eye movements are likely to be correlated with errors in their multi-attributes choices (ε). For these reasons, we expect our LIP model to better account for VA in the modelling of individuals' choices compared to the

[6] Rescaling procedure: First, for each respondent we computed the total fixation time per attribute across the 12 choice tasks. Second, we divided the total fixation times by 10,000, moving thus from measurements in milliseconds to measurements for 10 seconds. This transformation was done to avoid computational issues that occur when taking exp() of large numbers (e.g., exp(28)). Finally, for each attribute we computed a mean centred version of the rescaled total fixation times.

[7] Measurement error: Strictly speaking an eye-tracker only records location of point of gaze on the screen – this raw information is then converted into saccades (i.e., fast eyes movements) and fixations (i.e., period of time during which eyes are relatively still) using algorithms. Different algorithms would lead to a different number of fixations. Heterogeneous measures: A single fixation can vary widely in duration. In our study we observed fixation times ranging from 20 up to 280 milliseconds, suggesting that measures of attributes (non-)attendance based on fixation counts may actually obscure important between-attributes differences in the amount of visual attention. Understating the level of visual attention devoted to the attributes is likely to undermine the relationship between attention and preferences.

BIP model. This should be reflected by improvements in statistical fit (Log-likelihood; Bayesian Information Criterion, BIC)[8].

3.4 Comparison of willingness-to-pay (WTP) using different models

A crucial question is whether accounting for VA in different ways impacts on participants' WTP for changes in multi-attributes content. The computation of WTP slightly differs across models due to changes in the specification of the utility function (U)[9].

$$\mathrm{WTP_k} = -\frac{\partial \mathrm{U}/\partial \mathrm{X_k}}{\partial \mathrm{U}/\partial \mathrm{X_{cost}}} \qquad (15)$$

For the FIP-I and BIP models:

$$\mathrm{WTP_k} = -\frac{\beta_\mathrm{k}}{\beta_\mathrm{cost}} \qquad (16)$$

For the FIP-II model[10]:

$$\mathrm{WTP_k} = -\frac{\mu_\mathrm{k}}{\mu_\mathrm{cost}} \qquad (17)$$

For the LIP model[11]:

$$\mathrm{WTP_{nk}} = -\left(\frac{\beta_\mathrm{k}+\alpha_\mathrm{k}\mathrm{LIP_n}}{\beta_\mathrm{cost}+\alpha_\mathrm{cost}\mathrm{LIP_n}}\right) \qquad (18)$$

As indicated above, the homogeneous (FIP-I) and heterogeneous (FIP-II) version of the FIP model correspond respectively to the lower and upper bounds of variability in participants' choices. The WTP values obtained with FIP-II model are expected to systematically differ from those derived from FIP-I (i.e., all the WTP either increase or decrease, not necessarily to the same extent). As relaxing the hypothesis of full information processing can be seen as a partial but structured way of accounting for variability in participants' choices, we expect the WTP values associated with the BIP and LIP approaches to lie somewhere between those obtained with FIP-I and FIP-II.

[8] As the two models are not nested in each other, we cannot use log-likelihood ratio tests to compare the performance of the models.

[9] In practice the Delta method is commonly used to compute confidence intervals, however this method becomes impractical when the choice model includes randomness (e.g., random parameters in the mixed logit model). Standard bootstrapping provides an alternative approach, however this approach can be time consuming and impractical with simulated log-likelihood (e.g., FIP-II; LIP).

[10] In the FIP-II model, participants' preferences are assumed randomly distributed; a distribution of WTP values rather than a single average WTP value is thus derived. However to ease comparison with other choice models, we summarise information about the distribution of WTP values by taking its estimated mean.

[11] As indicated in Eq. 18, the marginal rate of substitution between two attributes is no longer a fixed quantity, but varies as a function of latent information processing ($\mathrm{LIP_n}$). To obtain a single measure of WTP that could be compared with those obtained from other choice models, we integrated for each attribute (k) the equation 18 over the LIP domain [-1;+1].

4. Results and discussion

Our CE contains 9,744 pieces of information (i.e., 58 participants × 12 choice tasks × 14 ROI) and overall 13.6% of the pieces of information were not consulted by the participants (Detailed results for the two experimental conditions can be found in supplementary material 1). Given that visual acuity is only sufficient to read small fonts (as used in our experiment) (Rayner, 1998), we can be fairly confident that these 13.6% pieces of information were not processed at all. This result is in line with Balcombe et al (2015) who found that "*most respondents visually attended most of the attributes most of the time*". In our CE, the last displayed attributes are more frequently ignored than the 1st displayed attributes (Figure 3). Also the pieces of information belonging to the 2nd (right) H&L choice option are systematically considered less. These descriptive results suggest the existence of both a *top-to-bottom* and *left-to-right* visual biases in participants' visual attention.

4.1 The Full Information Processing (FIP) model

The results of the FIP-I model are in line with *a priori* expectations regarding preferences for the attributes of H&L programmes (Table 3). All attributes, other than TIME and COST are associated with a positive effect on the utility of H&L programmes. For instance the probability of an H&L programme being selected significantly increases with improvements in the GOAL and DIABETES attributes. As expected, TIME and COST have a significant and negative effect, indicating that participants prefer both less expensive and less time consuming H&L programmes. In the FIP-I model, the ORDER parameter was positive and significant, indicating thus that choices from the 2nd experimental group (i.e., "reversed" order of the attributes) were more consistent than those from the 1st group (i.e., "initial" order). A log-likelihood ratio (LR) test indicated that the heteroscedastic version of the FIP-I model performed significantly better than its homoscedastic counterpart (LR test: Deviance = 13.4; DF = 1; P-val. < 0.001).

As predicted the FIP model allowing for heterogeneity in participants' preferences (FIP-II) provides a better account of multi-attributes choices (Log-likelihood: -471.4 vs. -524.1), which remain large even after accounting for increase in model complexity (BIC: 1,041.0 vs. 1,100.6)[12]. Most SD parameters are significant, indicating the presence of heterogeneity in preferences across participants (Table 3). Allowing for preferences heterogeneity does not noticeably alter the structure of preferences for the attributes.

[12] The heteroscedastic version of the FIP-II model does not perform better than its homoscedastic version (Log-likelihood: 471.8 vs. 471.4), and the ORDER parameter appeared to be non-significant (MLE (SE) = 0.353 (0.184); $P_{5\%}$ = 0.280). This suggests that most of the choices variability can be attributed to differences in individuals' preferences rather than changes of choice consistency. However, it is impossible to perfectly disentangle preferences heterogeneity from choice consistency (Hess and Rose 2012), and then in our FIP-II model the effect of ORDER parameter was also dependent on how we specified preferences heterogeneity.

4.2 The Binary Information Processing (BIP) model

The BIP model outperforms the FIP-I model (Log-likelihood: -508.4 vs. -524.1), indicating thus that part of the between-participants variability in choices can be linked to differences in VA (Table 3). However, as is shown in Figure 5, the statistical performance of the BIP model is sensitive to the choice of the cut-off point. For example, when specifying a 200 ms cut-off to differentiate non-attendance from attendance, 20.7% pieces of information is ignored by the participants and model log-likelihood was -508.35. At 400 ms we identify 39.1% pieces of information as being ignored by participants, and the model log-likelihood decreased to -535.12. Overall the model log-likelihood appears to falls quickly as the ANA proportion increases because of low specificity of the BIP approach (i.e., failing to recognise an attribute as being considered when it truly is). This result indicates that whilst the BIP model could be used to perform *data cleaning* (i.e., to exclude pieces of information with null fixation time) it is not suitable for carefully describing the relationship between VA and choice consideration.

4.3 The Latent Information Processing (LIP) model

Our LIP model outperforms the BIP model (Log-likelihood: -490.8 vs. -508.4) at the expense of seven additional parameters (i.e., interaction effects between LIP and preferences for the H&L attributes)[13].

Regarding the "information processing" component of this model (Table 4), all indicators of VA, other than PROG, are significant and positively related to the LIP variable, supporting our 1st assumption (**H1:** $\boldsymbol{\gamma}' \geq \mathbf{0}$). Participants with a higher LIP score spend more time fixating on the attributes. Assuming VA is related to information processing, this suggests that the LIP variable can effectively be interpreted as level of information uptake (i.e., "*how much individuals learn about the multi-attribute content*").

Turning to the multi-attributes choices (Table 3), the lower part of the Table shows the marginal change in preferences due to information uptake. Preferences for all the attributes but two are significantly influenced by changes in level of information uptake, providing some support for our 2nd assumption (**H2:** $\boldsymbol{\alpha}' \neq \mathbf{0}$).

Overall our assumptions regarding the effects of processing fluency on the relationship between VA and preferences are supported (**H3:** $\mathbf{r_{HBP}} = \mathbf{r_{DIABETES}} > \mathbf{r_{TIME}} = \mathbf{r_{COST}} > \mathbf{r_{WEIGHT}} \geq \mathbf{r_{PROG}} = \mathbf{r_{GOAL}}$). The preferences for the two imprecise and qualitative attributes (i.e., PROG; GOAL) are relatively immune to changes in level of information intake. The marginal effect of LIP on preferences for the PROG attribute represents 8.4% of the main effect

[13] The heteroscedastic version of the LIP model appeared to perform slightly better than the homoscedastic version (Log-likelihood: -489.9 vs. -490.8), but this difference did not reach significance (LR test: Deviance = 1.8; DF = 1; $P_{5\%} = 0.18$). Thus, we only report results for the homoscedastic LIP model. Results for the heteroscedastic model can be obtained from the corresponding author on request.

of the attribute on individuals' choices (0.049 / 0.582 = 0.084). The ratio for the GOAL attribute is 13.3%. However the interaction effects between LIP and the preferences are not significant. Next in the ranking come the TIME and COST attributes with ratios of 30.8% and 29% respectively. The two risk-related attributes, HBP and DIABETES, are even more impacted by changes in information uptake, with ratios of 58.1% and 49% respectively. A noticeable exception is the WEIGHT attribute which appears to be very sensitive to change in information uptake (Ratio = 80.8%), contrary to what we expected. This result indicates that the contribution of the WEIGHT attribute to the decision making mainly depends on amount of attention paid to the multi-attribute content of the H&L programmes. At low level of information uptake, the WEIGHT attribute will have no influence on individuals' choices.

These results show a significant contribution of VA to modelling multi attribute choice. Balcombe et al (2015) showed a weak contribution of measures of VA to the modelling of choices. One reason may be the difference in the context of the choice experiments. Balcombe et al (2015) asked participants to choose among different foods described by the traffic light system (i.e., colour coding of nutritional information). This experimental context may promote the use of visual decision heuristics (e.g., counting the number of green and red pieces of information) or at least to reinforce the effect of some visual biases (e.g., red pieces of information are more salient than yellow ones, and then have better "eyes-catching" properties, which in turn might bias respondents' decision making processes). Our context (health & lifestyle programmes) and presentation style (see example choice in Figure 2) may be more compatible with value-based choices and less prone to visual heuristics/biases, but this would require more research to establish. Another explanation may be the difference in the methods employed to account for VA – whilst Balcombe et al treat VA as a binary event (an attribute is considered as not being visually processed when it has been fixated less than twice), we maintain the continuous nature of visual attention (fixation times in milliseconds). Further, these two approaches differ in their representation of what we can learn from VA about individuals' choices. In Balcombe et al (2015), VA is limited to exploration of attributes non-attendance whilst in our study VA is a proxy for information processing to which ANA would be a particular case.

Our results have implications for CE practitioners. Finding that the underlying level of information processing has a significant effect on individuals' preferences contributes to the debate on the nature of stated preferences. While on one hand individuals' preferences are assumed to be well-defined and waiting to be measured, on the other hand there is ample evidence that preferences are ill-defined/malleable and "built on the fly" as participants progress through the CE (Bateman et al, 2008). If preferences were *a priori* defined, then changes in information processing should have little effect on their measurement. We find the opposite, suggesting that preferences are conditional on the context in which they have been elicited. Future research could use an eye-tracker to see how marginal changes in the multi-

Commented [nk7]: Reply #2

attribute content (e.g., adding/removing one attribute) impacts visual search behaviour and ultimately individuals' choices.

Further, we find that "easier to process" attributes are less influenced by changes in level of information processing than "difficult to process" attributes. Measurement of preferences for "difficult to process" attributes could be improved by (i) improving the user friendliness of the CE (i.e., devoting more attention to the quality of the interaction between participants and stimuli), for example this could be done by conducting a small quantitative pilot study with an eye-tracker, and (ii) prompting/forcing/nudging participants to pay more attention to difficult attributes. For example risk attributes could be located at the top of the choice options; participants could be reminded to carefully process the information; participants could be forced to look at the multi-attribute information for a minimum period of time (e.g., the next choice task would be automatically display after 40 seconds).

Our LIP model is based on structural equation modelling, and cannot be interpreted as causal (Bollen & Pearl, 2013). That is, we are learning how respondents pay attention to different types of information, and how that impacts choice. However, our approach does provide a structured framework to describe the expected relationship between visual information processing and individuals' choices. Fluency processing theory leads to testable hypotheses about the effect of visual information processing on preferences for the different attributes, depending on their nature. Further, it is difficult to investigate the causality of relationship between VA and individuals' choices, because they may influence each other in a feedback loop - *the longer you look at something, the more you like it; the more you like something, the longer you look at it*. Moving forward research could address this issue by explicitly modelling the evidence accumulation process during multi-attribute choices.

4.4 Comparison of willingness-to-pay (WTP) using different models

WTP values for the different models are shown in Table 5. Relaxing the assumption of preferences homogeneity impacts on WTP. We expected the BIP and LIP models to provide WTP values that fall between those obtained with the FIP-I and FIP-II models. This assumption was supported for our LIP model, but not for the BIP model. The BIP model tends to generate larger WTP values, suggesting not accounting for choice non-attendance produces biased welfare estimates. This finding is in line with previous ANA studies (Campbell, Hensher, and Scarpa, 2011). Further investigation revealed that the direction of the bias (i.e., underestimating vs overestimating WTP values) is mediated by a presentation effect (i.e., COST located at the top vs bottom of the choice options). In the "initial order" condition, the COST attribute was one of the least visually processed attribute (% visual ANA = 22.2); in the "reversed order" condition we find the opposite result, with "Cost" being the most frequently attended attribute (% visual ANA = 4.2). The BIP approach leads to higher and smaller WTP values for the "initial" and "reversed" condition respectively.

The more flexible LIP model leads to the expected correction of the WTP, with all values systematically smaller than those derived from the FIP-I model and larger than those obtained with the FIP-II model. This result suggests that between-subjects variability in VA would represent only one source of preferences heterogeneity.

Different WTP estimates can be explained by differences in how monetary valuations are corrected for using VA. By focusing only on cases where attributes are not visually processed, the BIP approach is limited to a partial correction of monetary values. Accounting for visual ANA shrinks the estimated preferences towards zero (in the "initial order" condition: $\beta_{\mathrm{COST}}^{\mathrm{FIP-I}} = -0.046$ vs. $\beta_{\mathrm{COST}}^{\mathrm{BIP}} = -0.039$), thereby generating higher WTP values (cost sensitivity being the denominator of the WTP formula). In principle it should be possible to correct WTP values for the opposite case of highly-attended attributes. Longer fixation time on an attribute can be an indication of stronger preference for the attribute. Previous studies showed that individuals tend to spend more time looking at what they like (Milosavljevic et al, 2012; Shimojo et al, 2003). As our LIP model assumes a continuous effect of information processing on multi-attribute choices, it is able to correct monetary valuations for both extremely low and high level of information processing.

Further, when specifying a linear additive utility function, WTP values are computed by dividing estimated preferences by the preference for the COST dimension; resulting WTP values are thus sensitive to changes in the denominator. With the BIP approach, WTP values depend on how VA is translated into measures of ANA, questioning thus the validity of this approach (as at the moment there is no golden rule to convert fixation times or counts into meaningful measures of information processing). In this regard it may be hypothesised that our proposed LIP model provides a more sensible correction of the WTP values. A simulation study comparing different decision rules would be welcome.

5. Concluding remarks

Choice experiments are commonly used in economics to value public or quasi-public goods. When measuring preferences, unwarranted assumptions are often made about how individuals process information about the attributes of the products or services. The classical full information processing (FIP) model assumes individuals are able to all process information without restrictions. We develop a new framework in which information processing is treated as a latent (unobservable) process. Testing our approach by combining CE and visual attention (VA) data gathered from eye-tracking, we show that treating information processing as a latent process outperforms the FIP model. We show that the relationship between VA and individuals' preferences depends on the type of product attribute. More specifically, preferences for "easier to process" attributes appear to be less influenced by changes in underlying level of VA than "harder to process" attributes. In turn this impacts on willingness-to-pay estimates. Our results have implications for CE designers.

Commented [nk8]: Reply #7.1

More time should be spent getting subjects to understand more complicated attributes of the CE. Our results are likely not limited to the field of experimental choices (stated preferences), and can also inform analyses of actual choices (revealed preferences).

6. Acknowledgments

The University of Aberdeen (UoA) and the Chief Scientist Office (CSO) of the Scottish Government Health and Social Care Directorates fund the Health Economics Research Unit (HERU). We thank Pavlos Topalidis for help with data collection and thank all participants for taking part. The views expressed in this paper are those of the authors and not those of the funding bodies.

Bibliography

Ambinder, M. S., and D. J. Simons. 2010. "Individual Differences in Attention Capture." *Journal of Vision* 6 (6): 589–589. doi:10.1167/6.6.589.

Balcombe, Kelvin, Iain Fraser, and Eugene McSorley. 2015. "Visual Attention and Attribute Attendance in Multi-Attribute Choice Experiments: Discrete Choice Experiments and Eye-Tracking." *Journal of Applied Econometrics* 30 (3): 447–67. doi:10.1002/jae.2383.

Bateman, Ian J., Diane Burgess, W. George Hutchinson, and David I. Matthews. 2008. "Learning Design Contingent Valuation (LDCV): NOAA Guidelines, Preference Learning and Coherent Arbitrariness." *Journal of Environmental Economics and Management* 55 (2): 127–41. doi:10.1016/j.jeem.2007.08.003.

Ben-Akiva, Moshe, Daniel McFadden, Kenneth Train, Joan Walker, Chandra Bhat, Michel Bierlaire, Denis Bolduc, et al. 2002. "Hybrid Choice Models: Progress and Challenges." *Marketing Letters* 13 (3): 163–75.

Bollen, K. A., and J. Pearl. 2013. "Eight Myths about Causality and Structural Equation Models." In *Handbook of Causal Analysis for Social Research*, Springer, 301–328.

Campbell, Danny, David A. Hensher, and Riccardo Scarpa. 2011. "Non-Attendance to Attributes in Environmental Choice Analysis: A Latent Class Specification." *Journal of Environmental Planning and Management* 54 (8): 1061–76. doi:10.1080/09640568.2010.549367.

Carlsson, Fredrik, Mitesh Kataria, and Elina Lampi. 2010. "Dealing with Ignored Attributes in Choice Experiments on Valuation of Sweden's Environmental Quality Objectives." *Environmental and Resource Economics* 47 (1): 65–89. doi:10.1007/s10640-010-9365-6.

Caussade, Sebastián, Juan de Dios Ortúzar, Luis I. Rizzi, and David A. Hensher. 2005. "Assessing the Influence of Design Dimensions on Stated Choice Experiment Estimates." *Transportation Research Part B: Methodological* 39 (7): 621–40. doi:10.1016/j.trb.2004.07.006.

Cavanagh, James F., Thomas V. Wiecki, Angad Kochar, and Michael J. Frank. 2014. "Eye Tracking and Pupillometry Are Indicators of Dissociable Latent Decision Processes." *Journal of Experimental Psychology: General* 143 (4): 1476–88. doi:10.1037/a0035813.

Czajkowski, M., N. Hanley, and J. LaRiviere. 2015. "The Effects of Experience on Preferences: Theory and Empirics for Environmental Public Goods." *American Journal of Agricultural Economics* 97 (1): 333–51. doi:10.1093/ajae/aau087.

Czajkowski, Mikołaj, Nick Hanley, and Jacob LaRiviere. 2016. "Controlling for the Effects of Information in a Public Goods Discrete Choice Model." *Environmental and Resource Economics* 63 (3): 523–44. doi:10.1007/s10640-014-9847-z.

de Bekker-Grob, Esther W., Mandy Ryan, and Karen Gerard. 2012. "Discrete Choice Experiments in Health Economics: A Review of the Literature." *Health Economics* 21 (2): 145–72. doi:10.1002/hec.1697.

DeShazo, J.R., and German Fermo. 2002. "Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency." *Journal of Environmental Economics and Management* 44 (1): 123–43. doi:10.1006/jeem.2001.1199.

Duchowski, Andrew T. 2007. *Eye Tracking Methodology: Theory and Practice*. 2nd ed. London: Springer.

Fagerlin, A., B. J. Zikmund-Fisher, and P. A. Ubel. 2011. "Helping Patients Decide: Ten Steps to Better Risk Communication." *JNCI Journal of the National Cancer Institute* 103 (19): 1436–43. doi:10.1093/jnci/djr318.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg. 2006. "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model." *American Economic Review* 96 (4): 1043–68. doi:10.1257/aer.96.4.1043.

Gaglio, Bridget, Russell E. Glasgow, and Sheana S. Bull. 2012. "Do Patient Preferences for Health Information Vary by Health Literacy or Numeracy? A Qualitative Assessment." *Journal of Health Communication* 17 (sup3): 109–21. doi:10.1080/10810730.2012.712616.

Hensher, David A., and William H. Greene. 2010. “Non-Attendance and Dual Processing of Common-Metric Attributes in Choice Analysis: A Latent Class Specification.” *Empirical Economics* 39 (2): 413–26. doi:10.1007/s00181-009-0310-x.

Hess, Stephane, and David A. Hensher. 2013. “Making Use of Respondent Reported Processing Information to Understand Attribute Importance: A Latent Variable Scaling Approach.” *Transportation* 40 (2): 397–412. doi:10.1007/s11116-012-9420-y.

Hess, Stephane, and John M. Rose. 2012. “Can Scale and Coefficient Heterogeneity Be Separated in Random Coefficients Models?” *Transportation* 39 (6): 1225–39. doi:10.1007/s11116-012-9394-9.

Hole, Arne Risa. 2011. “A Discrete Choice Model with Endogenous Attribute Attendance.” *Economics Letters* 110 (3): 203–5. doi:10.1016/j.econlet.2010.11.033.

Hole, Arne Risa, Julie Riise Kolstad, and Dorte Gyrd-Hansen. 2013. “Inferred vs. Stated Attribute Non-Attendance in Choice Experiments: A Study of Doctors’ Prescription Behaviour.” *Journal of Economic Behavior & Organization* 96: 21–31. doi:10.1016/j.jebo.2013.09.009.

Holmqvist, Kenneth, ed. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford ; New York: Oxford University Press.

Hoyos, David. 2010. “The State of the Art of Environmental Valuation with Discrete Choice Experiments.” *Ecological Economics* 69 (8): 1595–1603. doi:10.1016/j.ecolecon.2010.04.011.

Jasper, John D., Chandrima Bhattacharya, and Ryan Corser. 2016. “Numeracy Predicts More Effortful and Elaborative Search Strategies in a Complex Risky Choice Context: A Process-Tracing Approach: Numeracy and Process Tracing.” *Journal of Behavioral Decision Making*, n/a-n/a. doi:10.1002/bdm.1934.

Joško Brakus, J., Bernd H. Schmitt, and Shi Zhang. 2014. “Experiential Product Attributes and Preferences for New Products: The Role of Processing Fluency.” *Journal of Business Research* 67 (11): 2291–98. doi:10.1016/j.jbusres.2014.06.017.

Keller, Carmen, Christina Kreuzmair, Rebecca Leins-Hess, and Michael Siegrist. 2014. “Numeric and Graphic Risk Information Processing of High and Low Numerates in the Intuitive and Deliberative Decision Modes: An Eye-Tracker Study.” *Judgment and Decision Making* 9 (5): 420–432.

Keller, Carmen, Michael Siegrist, and Vivianne Visschers. 2009. “Effect of Risk Ladder Format on Risk Perception in High- and Low-Numerate Individuals.” *Risk Analysis* 29 (9): 1255–64. doi:10.1111/j.1539-6924.2009.01261.x.

Keller, Punam Anand, and Ann L. McGill. 1994. “Differences in the Relative Influence of Product Attributes under Alternative Processing Conditions: Attribute Importance versus Attribute Ease of Imagability.” *Journal of Consumer Psychology* 3 (1): 29–49. doi:10.1016/S1057-7408(08)80027-7.

Kjaer, Trine, Mickael Bech, Dorte Gyrd-Hansen, and Kristian Hart-Hansen. 2006. “Ordering Effect and Price Sensitivity in Discrete Choice Experiments: Need We Worry?” *Health Economics* 15 (11): 1217–28. doi:10.1002/hec.1117.

Krajbich, I., and A. Rangel. 2011. “Multialternative Drift-Diffusion Model Predicts the Relationship between Visual Fixations and Choice in Value-Based Decisions.” *Proceedings of the National Academy of Sciences* 108 (33): 13852–57. doi:10.1073/pnas.1101328108.

Lagarde, Mylene, and Duane Blaauw. 2009. “A Review of the Application and Contribution of Discrete Choice Experiments to Inform Human Resources Policy Interventions.” *Human Resources for Health* 7 (1): 62. doi:10.1186/1478-4491-7-62.

Lancaster, Kelvin J. 1966. “A New Approach to Consumer Theory.” *Journal of Political Economy* 74 (2): 132–57.

LaRiviere, Jacob, Mikołaj Czajkowski, Nick Hanley, Margrethe Aanesen, Jannike Falk-Petersen, and Dugald Tinch. 2014. “The Value of Familiarity: Effects of Knowledge and Objective Signals on Willingness to Pay for a Public Good.” *Journal of Environmental Economics and Management* 68 (2): 376–89. doi:10.1016/j.jeem.2014.07.004.

Liechty, John, Rik Pieters, and Michel Wedel. 2003. “Global and Local Covert Visual Attention: Evidence from a Bayesian Hidden Markov Model.” *Psychometrika* 68 (4): 519–41. doi:10.1007/BF02295608.

Luce, R. Duncan. 2005. *Individual Choice Behavior: A Theoretical Analysis*. Mineola, N.Y: Dover Publications.

Manski, Charles F. 1977. “The Structure of Random Utility Models.” *Theory and Decision* 8 (3): 229–54. doi:10.1007/BF00133443.

McFadden, Daniel. 1974. “Conditional Logit Analysis of Qualitative Choice Behavior.” In *Frontier in Econometrics*, 105–42. New York: Academic Press.

McFadden, Daniel, and Kenneth Train. 2000. “Mixed MNL Models for Discrete Response.” *Journal of Applied Econometrics* 15 (5): 447–70. doi:10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1.

Milosavljevic, Milica, Vidhya Navalpakkam, Christof Koch, and Antonio Rangel. 2012. “Relative Visual Saliency Differences Induce Sizable Bias in Consumer Choice.” *Journal of Consumer Psychology* 22 (1): 67–74. doi:10.1016/j.jcps.2011.10.002.

Nunez, Michael D. 2015. “Individual Differences in Attention Influence Perceptual Decision Making.” *Frontiers in Psychology* 6. doi:10.3389/fpsyg.2015.00018.

Osimani, Barbara. 2012. “Risk Information Processing and Rational Ignoring in the Health Context.” *The Journal of Socio-Economics* 41 (2): 169–79. doi:10.1016/j.socec.2011.10.009.

Peters, E., P. S. Hart, M. Tusler, and L. Fraenkel. 2014. “Numbers Matter to Informed Patient Choices: A Randomized Design across Age and Numeracy Levels.” *Medical Decision Making* 34 (4): 430–42. doi:10.1177/0272989X13511705.

Rayner, Keith. 1998. “Eye Movements in Reading and Information Processing: 20 Years of Research.” *Psychological Bulletin* 124 (3): 372–422. doi:10.1037/0033-2909.124.3.372.

Reutskaja, Elena, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel. 2011. “Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study.” *American Economic Review* 101 (2): 900–926. doi:10.1257/aer.101.2.900.

Ryan, Mandy, Deokhee Yi, Alison Avenell, Flora Douglas, Lorna Aucott, Edwin van Teijlingen, and Luke Vale. 2015. “Gaining Pounds by Losing Pounds: Preferences for Lifestyle Interventions to Reduce Obesity.” *Health Economics, Policy and Law* 10 (2): 161–82. doi:10.1017/S1744133114000413.

Scarpa, R., S. Notaro, J. Louviere, and R. Raffaelli. 2011. “Exploring Scale Effects of Best/Worst Rank Ordered Choice Data to Estimate Benefits of Tourism in Alpine Grazing Commons.” *American Journal of Agricultural Economics* 93 (3): 813–28. doi:10.1093/ajae/aaq174.

Shah, Anuj K., and Daniel M Oppenheimer. 2007. “Easy Does It: The Role of Fluency in Cue Weighting.” *Judgment and Decision Making* 2 (6): 371–79.

Shimojo, Shinsuke, Claudiu Simion, Eiko Shimojo, and Christian Scheier. 2003. “Gaze Bias Both Reflects and Influences Preference.” *Nature Neuroscience* 6 (12): 1317–22. doi:10.1038/nn1150.

Sims, Christopher A. 2006. “Rational Inattention: Beyond the Linear-Quadratic Case.” *American Economic Review* 96 (2): 158–63. doi:10.1257/000282806777212431.

Thurstone, L. L. 1927. “A Law of Comparative Judgment.” *Psychological Review* 34 (4): 273–86. doi:10.1037/h0070288.

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation*. 2nd ed. Cambridge ; New York: Cambridge University Press.

Valuch, Christian, Lena S. PflÃ¼ger, Bernard Wallner, Bruno Laeng, and Ulrich Ansorge. 2015. “Using Eye Tracking to Test for Individual Differences in Attention to Attractive Faces.” *Frontiers in Psychology* 6 (February). doi:10.3389/fpsyg.2015.00042.

Yang, Cathy L, Oliver Toubia, and Martijn G de Jong. 2013. “Incentive-Compatible Preference Measurement with Boundedly Rational Consumers.”