Visual Attention and Attribute Attendance in Multi-Attribute Choice Experiments.

Kelvin Balcombe, University of Reading
Iain Fraser, University of Kent/La Trobe University
Eugene McSorley, University of Reading

Contributed Paper prepared for presentation at the 87th Annual Conference of the Agricultural Economics Society, University of Warwick, United Kingdom
8 - 10 April 2013

Address for correspondence: School of Economics, University of Kent Canterbury, Kent, CT2 7NP, UK, Tel: +44 (0)1227 823513, i.m.fraser@kent.ac.uk

Abstract

Decision strategies in multi-attribute Choice Experiments are investigated using eye-tracking. The visual attention towards, and attendance of, attributes is examined. Stated attendance is found to diverge substantively from visual attendance of attributes. However, stated and visual attendance are shown to be informative, non-overlapping sources of information about respondent utility functions when incorporated into model estimation. Eye-tracking also reveals systematic nonattendance of attributes only by a minority of respondents. Most respondents visually attend most attributes most of the time. We find no compelling evidence that the level of attention is related to respondent certainty, or that higher or lower value attributes receive more or less attention.

Key Words: Choice Experiment, Attribute Nonattendance, Eye-tracking, Random Utility Models.

JEL: C1,C35,D12

Visual Attention and Attribute Attendance in Multi-Attribute Choice Experiments.

1. Introduction

The applied economics literature has witnessed widespread application of multi-attribute Choice Experiments (CEs) as it has become the preferred approach to stated preference research. CEs present survey participants with a set of attributes of varying levels that are used to describe a good which we are interested in valuing. By varying the level of the attributes across several choice situations researchers examine how choices change. CEs generally suppose that stated choices are the outcome of interrogation by the respondent of their own (random) utility function. Random Utility Models (RUMs) (i.e., McFadden 2001) provide the key theoretical underpinning for CEs. As such the utility function is assumed to exist independently of the experiment, and the respondent must be willing and able to give responses consistent with that utility function. The RUM model provides the theoretical justification for the estimation of willingness-to-pay estimates for attributes and welfare measures from CEs (McFadden and Train, 2000).

While many economists accept that RUMs offer a reasonable approximation of respondent behaviour, fewer economists would argue that all respondents within a CE always act in strict accordance with a RUM. Indeed, there is a growing literature that questions the validity of RUM (e.g., Kahneman, 2003, and DellaVigna, 2009). As a result the CE literature contains examples whereby the utility function is treated as a temporary construct, which is shaped by the CE. For example, Hensher (2006) posit that CEs of differing complexity lead to information processing strategies that result in alternative utility functions being used in different experiments. Strands of literature employ models of stochastic choice that are openly not RUMs. See Marley and Louviere (2005) for some examples. Some non-RUMs posit that people adopt choice strategies that depend on many factors, including the complexity of the choice tasks. The outcome of these strategies may be choices that are inconsistent with a RUM, and undermine the internal and external validity of the CE. This necessitates an investigation of respondent behaviour that draws upon information other than that which can be inferred by stated responses. Eye-tracking provides one of the most powerful means by which the processes driving individuals' choices can be uncovered.

One potential strategy that respondents might adopt is to ignore certain attributes of the good in question when making a choice. Within the CE literature this has been labelled as 'attribute nonattendance' (ANA). A growing number of studies have examined ANA.(e.g., Hensher et al., 2005, Scarpa et al., 2010, and Balcombe et al., 2011). That an attribute has zero or negligible utility is not, in itself, at odds with a RUM. However, nonattendance of a subset of attributes could signal that an individual is adopting choice strategies inconsistent with a RUM (e.g. lexicographic, (Tversky, 1972) or random regret (Chorus et al., 2008)). Therefore, knowing whether or not respondents are attending attributes offers information about respondents utility functions within a RUM, but also offers an insight into whether they are adopting choice strategies that are consistent with a RUM.

To date the literature has investigated ANA by either inferring ANA based on respondent choices or by asking them *ex post* de-briefing questions about their attendance of attributes. In this paper we employ eye-tracking technology to monitor the way in which survey participants engage with a CE survey instrument. Supplementing *ex post* debriefing questions about ANA with other objective measures of ANA is crucial, because *ex post* debriefing responses are not always an accurate means to recover information about actual behaviour. The use of eye-tracking technology allows us to examine if eye fixations are consistent with respondent reporting of ANA, and assess if either provides useful information about the choice strategies of respondents. Moreover, we can assess whether 'higher value' attributes are given higher visual attention and whether respondent uncertainty, as indicated by scale heterogeneity, is explained by total visual attention paid to choice tasks.

The use of eye-tracking is relatively new to economics, although it has a longer history within marketing and psychology. For example, it has been found that the level of attention to a given brand on a shelf, measured with use of eye-tracking, is related to subsequent purchase choice (e.g., Wedel and Pieters, 2007, and Aribarg et al., 2010). Within economics, eye-tracking is associated with neuroeconomics (e.g., Caplin and Dean, 2008). For example, eye-tracking has been used to experimentally examine visual search and how this relates to decision making (e.g., Knoepfle et al., 2009 and Reutskaja et al., 2011), and in sender-receiver games it has been found that pupil dilation is correlated with deception (Wang et al., 2010). More recently, Caplin et al. (2011) suggest that eye-tracking could be used in combination with choice process experiments to better understand how economic agents undertake search activities. However, the use of eye-tracking in hypothetical multi-attribute CEs has not been explored.

For our empirical work we employ a CE survey instrument that has previously been tested in Balcombe et al. (2010). This survey instrument was designed to examine consumer understanding of the United Kingdom nutrient content food label, the Traffic Light System (TLS), that was used to indicate nutrient levels on processed food. The design of this CE is standard in terms of the number of choices and attributes. Earlier experience using this instrument suggested that respondents understood the TLS in terms of making choices and how it related to the choice tasks. However, resulting willingness-to-pay estimates for reductions in nutrients are considered high and this is frequently taken as a possible sign of ANA. As such, we considered this survey instrument to be suitable to employ to examine visual ANA using eye-tracking.

We proceed as follows. In Section 2 we develop a detailed definition of ANA. This will inform both the way in which we interpret our eye-tracking results. In section 3 we introduce and explain eye-tracking. The econometric model is outlined in Section 4. This is followed by a description of the eye-tracking experiment in section 5 and in Section 6 we report and discuss the results of our experiment. Section 7 concludes.

2. Defining ANA

When defining and explaining ANA for a multi-attribute CE, it is helpful to make a distinction between measures of attendance and attention. A respondent may have paid attention to an attribute but still not be considered to have attended the attribute, either because it has played no role in their choices or because they have only paid attention to a subset of the relevant information. As we shall define it, attendance is a discrete measure. Respondents will either be considered to have attended an attribute or not. In contrast, attention, is a continuous measure of the degree to which a respondent evaluates attribute levels.

ANA has generally been defined to mean that the variation in the levels of given attribute(s) has no influence over respondents choices. However, it has also been used implicitly in a similar but distinct way. That is, ANA is also taken to imply that information about the attribute levels has not been 'processed'. In order to be clear in our empirical analysis, we employ the following definitions of ANA within the paper:

Definition 1. Information ANA - Not all the information provided about an attribute's levels is processed during the CE.

Definition 2. Choice ANA - The levels of an attribute presented to respondents have played no role in determining respondent choices within a CE.

Definition 3. Stated ANA - The individual has stated *ex post* that they have ignored an attribute when completing the CE.

Each definition can exist in a serial sense or be specific to a subset of choices. Serial ANA implies ANA over all choice tasks. In the majority of studies, stated ANA has been elicited by a question after the completion of all choice tasks. Therefore, these studies have arguably asked the respondent to declare serial ANA.

Within RUMs, choice ANA could be interpreted as the respondent having zero marginal utility for the attribute in question. In related literature, Marley et al. (2008) refer to what they call the importance of an attribute and the utility of an attribute-level. They observe that it is difficult to separate these two factors and so do not refer to zero marginal utility but instead use the phrase 'zero marginal impact'. Regardless of whether we take ANA to imply zero marginal utility or impact, establishing choice ANA can be viewed as an important characterisation about the respondents' utility function.

Furthermore, ANA may be a sign that a respondent is acting in a way that is inconsistent with a RUM. This is particularly so when a respondent states that they ignored the price attribute, and/or when they ignore multiple other attributes. Consequently, there is considerable discussion about why ANA occurs. For example, Hensher et al. (2005) argue that in CEs an individual can employ various information processing strategies and that ANA could arise in a number of ways including: (i) a coping strategy to deal with task complexity; (ii) the cost from evaluating attribute(s); or, (iii) an attribute 'truly' not influencing choice.

So is stated ANA a good (or perfect) indicator of choice and/or information ANA? This has been investigated in two ways. First, by comparing inferred ANA (i.e. that attribute levels appear to have played no role in determining an individuals choice, as inferred by their choice(s) within the CE) with stated ANA. Second, by testing the restriction that stated ANA equals choice ANA within the estimated utility function. Most studies find that stated ANA cannot reliably be interpreted as choice ANA. Perhaps stated ANA indicates that certain attributes played a small or marginal role, or perhaps it means something entirely different. For example, someone may say they ignored the salt content in food, which could mean that they are signalling that while they realised that too much salt is bad for them, they chose high salt foods anyway.

Given this lack of clarity in the literature we propose that other ways are needed to investigate ANA. Since CEs are nearly always presented to respondents visually, respondents may not look at some of the information presented to them, or may not have looked long enough for the information to have been processed. Therefore, to definitions 1)-3) we add an additional type of ANA:

Definition 4. Visual ANA - Some (or all) of the information about the attribute levels has been visually ignored.

Visual ANA can be investigated empirically, since there is technology to measure visual attention, and there is an associated body of knowledge that indicates that eye-tracking is able to discern whether respondents have looked at information, and whether their eyes have 'fixed' for long enough upon that information for it to have been used.

3. Eye-tracking: A Brief Introduction

An individual will tend to move their eyes when confronted with visual stimuli. The physical reason why this happens is because sharpness across the retina declines quickly with distance away from the fovea: the part of the eye responsible for processing fine-grained, detailed visual information. Only about 2% of the visual field is projected onto the fovea which means that in order to subject stimuli to scrutiny the eye have to be moved. Therefore, examining eye-movements can be potentially important in understanding information acquisition (Rayner, 2009, McSorley and McCloy, 2009, and McSorley et al., 2009).

Eye movements are not smooth, and are composed of two separate elements: fixations and saccades. Fixations describe movements when the eye is 'relatively' still. In general viewing fixations have durations of between 200-500 milliseconds in which a contiguous area is projected onto the fovea allowing detailed visual processing. In contrast, saccades are very rapid movements shifting gaze to areas of interest and taking as little as 20-40 milliseconds. This type of movement helps project specific locations of a scene onto the fovea.

In combination with understanding eye movements themselves, eye-tracking research also aims to understand how the brain deals with information received. This information, which is transmitted via the optic nerve, is greater than the brain can deal with, as a result humans have developed various attentional mechanisms that aid in the selection of a subset of relevant information that is subject to enhanced processing. This means that the brain is simultaneously enhancing and suppressing information.

In normal viewing situations attention and eye movements are intimately linked and move in tandem to the same visual location (Deubel and Schneider, 1996, and Findlay, 2009). This comes from evidence examining the close correspondence between eye movements and higher-order cognitive processes (e.g., Rayner, 2009). As such eye-tracking research has provided insights into the control of visual attention (Awh et. al., 2006, Findlay, 2009, and Theeuwes, 2010).

In practice, eye-tracking research looks for patterns based on fixations and saccades. Eye-trackers record patterns of these movements and pauses, while people view a visual stimulus. These patterns are then collected together in what is referred to as a scanpath. This scanpath provides spatial-temporal data on spatial distribution of attention across the visual stimulus. Therefore, eye fixation is in principle a good indicator of visual attention because (i) acuity deteriorates rapidly outside the fovea; (ii) little visual information can be obtained during saccades (Matin, 1974); and (iii) fixation and attention are naturally yoked.

The general consensus in the economics literature is that initial fixation activity is random and as such does not generate what Knoepfle et al. (2009) refer to as information lookup. It is only when a respondent fixates more than once (referred to as a refixation) that an assessment of (relative) value occurs.

Finally, we also note that there are methodological variants within the eye-tracking literature. In particular, Franco-Watkins and Johnson (2011,a and b) have introduced a variation on the typical form of eye-tracking called the decision moving window. This technique is related to the Flashlight method developed by Shulte-Mecklenbeck et al. (2011). These alternative forms of eye-tracking method highlight that there are a wide range of options for the use of eye-tracking to enhance economics choice based research.

4. The Econometric Specification

This section outlines the econometric methods that we employ to estimate the utility functions from the CE. We will start with a brief statement of the standard 'mixed logit' then generalise this model to allow for sequential and individual scale heteroscedasticity and marginal utilities that are dependent on visual and stated nonattendance data.

4.1.1. The Standard 'Mixed Logit'

In order to describe our model we shall first start by describing the standard mixed logit. The utility that the *jth* individual receives from the *ith* option in the *sth* choice set is assumed to be of the form

$$\dot{U}_{ijs} = \dot{x}'_{ijs}\dot{\beta}_j + e_{ijs} \tag{1}$$

where $\dot{x}_{ijs}$ denotes the $k \times 1$ vector of attribute levels presented to the *jth* individual $(j = 1, .., J)$ in the *ith* option $(i = 1, .., I)$ of the *sth* choice set $(s = 1, .., S)$. The error e_{ijs} is 'extreme value' (Gumbel) distributed, is independent of $\dot{x}_{ijs}$, and is uncorrelated across individuals or across choices. $\dot{\beta}_j$ is a $(k \times 1)$ vector describing the preferences of the *jth* individual and obeys

$$\dot{\beta}_j = \alpha + u_j \tag{2}$$

where α is the mean of $\dot{\beta}_j$ and u_j is an independently and identically normally distributed vector with variance-covariance matrix Ω that is not constrained to be diagonal. The errors $\{u_j\}$ are assumed to be uncorrelated across individuals. In what follows we will continue to refer to the values of $\dot{\beta}_j$ as the marginal utilities even when the utility function (1) is scaled.

4.1.2. Scale Heteroscedasticity

There has been considerable interest in scale heteroscedasticity (e.g. Fiebig et al., 2010). Evidence from the CE literature (e.g. Savage and Waldman, 2008) suggests that there may be learning and/or fatigue by respondents through the course of a CE, that may be captured by sequence scale heteroscedasticity in the Gumbel error. Since eye-tracking potentially provides information about levels of attention, as well as changes in attention, we extend our model to allow for scale heteroscedasticity, to see whether attention as measured by eye-tracking fixations improves model performance.

If the variance of the Gumbel error $\{\sigma\}$ is specified as dependent on s, the utility function becomes

$$\dot{U}_{ijs} = \dot{x}'_{ijs}\dot{\beta}_j + \sigma_{js}e_{ijs} \tag{3}$$

In principle, a variance can be independently estimated for each s. However, this approach ignores the likely smoothness in the function σ_{js} within the choice sets s. Therefore, it is advantageous to put a functional form on σ_{js}. We specify the following functional form for the scale variance

$$\sigma_{js} = e^{\left(-\phi_1(\omega_s - \bar{t}_s) - \phi_2\left(\sin(\omega_s\pi) - \overline{\sin(\omega_s\pi)}\right) - \phi_3\left(\sin(\omega_s 2\pi) - \overline{\sin(\omega_s 2\pi)}\right) - \phi_4 z_{js}\right)} \tag{4}$$

where $\omega_s = \left(\frac{s-1}{S-1}\right)$, and the parameters $\phi = \{\phi_v\}_{v=1}^{4}$ are to be estimated. In (4) the first term is linear so that if the variance increases or decreases throughout the experiments $\phi_1 \neq 0$. The second is a sinusoidal function that peaks in the middle of responses provided. This type of variance behaviour happens if respondents learn within the first half of the CE, but then become fatigued in the second half of the CE. The third term gives further flexibility to the relationship so that the maximum or minimum of the scale variance can be at other points in the first or second half of the choice sets given to individuals.

By allowing the first three coefficients in (4)($\{\phi_v\}_{v=1}^{3}$) to be estimated, we can flexibly approximate a range of behaviours in terms of learning and fatigue. We subtract the means of the trends and sinusoidal function so that the average variance (over s) is approximately one. Thus, the models with heteroscedasticity are more comparable with a model where $\sigma_{js} = 1$, for s $=1,....S$.

The last term z_{js} is the log of fixation duration by the *jth* individual on the *sth* choice card (normalised to have mean zero). Thus, if ϕ_4 is positive, then those individuals who dwelt on a choice card for longer tended to have a much lower variance attached to the Gumbel error, and, in a sense, are more certain about their choice.

4.1.3. Using Attribute Nonattendance Data

In (3) we defined utility as

$$\dot{U}_{ijs} = \dot{x}'_{ijs}\dot{\beta}_j + \sigma_{js}e_{ijs} \tag{5}$$

We now assume that we have information about attendance that we use to modify the distribution of the original marginal utilities $\left\{\dot{\beta}_j\right\}$. In order to do this we introduce a matrix Λ_j and write the utility function as:

$$\dot{U}_{ijs} = \dot{x}'_{ijs}\Lambda_j\beta_j + \sigma_{js}e_{ijs} \tag{6}$$

where

$$\beta_j \sim N(\alpha, \Omega) \tag{7}$$

Thus, $\{\beta_j\}$ has the normal distribution $N(\alpha, \Omega)$ previously assigned to the marginal utilities $\left\{\dot{\beta}_j\right\}$. This is equivalent to assuming that an individual's marginal utilities (i.e., the marginal of $\dot{U}_{ijs}$ with respect to $\dot{x}'_{ijs}$) have the distribution

$$\dot{\beta}_j \sim N\left(\Lambda_j\alpha, \Lambda_j\Omega\Lambda'_j\right) \tag{8}$$

By model design the matrix $\Lambda_j = diag(\lambda_{j1}, ...\lambda_{jK})$ is constructed from the nonattendance data and has the elements

$$\lambda_{jk} = \prod_{c=1}^{C}(1 - \delta_{cjk} + \tau_c\delta_{cjk}) \tag{9}$$

and δ_{cjk} is an indicator variable that takes the value 1 if the jth person is classified as a nonattender of the kth attribute according to criteria c. We assume that the parameter τ_c is bounded between the unit interval [0,1]. In our empirical examples we consider two criteria: stated ($c = 1$) and visual ($c = 2$) nonattendance. In both cases the lower bound $\tau_c = 0$ means that a nonattender ($\delta_{cjk} = 1$) has zero marginal utility for an attribute that they do not attend and at $\tau_c = 1$ there is no difference between the distributions of the marginal utility of the attender and nonattender. In general, the lower the value of τ_c the greater the 'shrinkage' of marginal utility towards zero. This has similarities to the approach taken by Scarpa et. al. (2010), though here we allow for any value of τ_c between 0 and 1.

In the case where an individual is defined to non-attend simultaneously with respect to both criteria, then the magnitude of the shrinkage of marginal utility for the attribute concerned will be $\tau_1 \times \tau_2$. This approach can be extended to include attribute specific shrinkage parameters. However, for identification of these parameters there would need to be relatively large samples with reasonably large numbers of nonattenders with respect to each of the attributes.

The approach to ANA which we are employing is, therefore, one in which each individual is characterised as a serial attender or nonattender throughout the CE. In principle this approach could be adapted so that the indicator variables $\{\delta_{cjk}\}$ become choice specific. While this approach is more general from an econometric perspective, varying marginal utilities over the choice sets has no strong theoretical motivation. Therefore, we do not explore the choice specific approach here.

Finally, we note that there are potential alternatives to the treatment of serial nonattendance data. For example, we could in principle employ the nonattendance indicators as covariates for the latent marginal utilities, while leaving the variance of the marginal utilities unchanged. However, the assumption that the variance for nonattenders would also diminish is more likely and that $\tau_c = 0$ represents a special case of a more general model, makes the approach taken here more attractive. The covariate approach is also problematic for various reasons. For example, take an attribute that has marginal utilities that are positive for part of the population and negative for another part with a mean of zero. The covariate treatment of nonattendance for that attribute must be positive or negative yet moving the mean of the distribution for nonattenders to the left or right makes little sense. This runs counter to the reason for using nonattendance information, which is based on the idea that nonattenders will have marginal utilities closer to zero. The covariate approach also becomes even more problematic when considering multiple ANA measures because the aggregate effects can imply large negative or positive estimates of marginal utilities for multiple nonattenders.

4.2. Estimation

This model is reasonably simple to estimate using Bayesian methods, since it can be specified in a similar way to the standard Mixed Logit, with the normal latent variables being multiplied by the shrinkage terms. In (7) we assumed that $\beta_j \sim N(\alpha, \Omega)$. By defining

$$x'_{ijs} = \sigma_{js}^{-1} \dot{x}'_{ijs} \Lambda_j \text{ and } U_{ijs} = \dot{U}_{ijs} \sigma^{-1} \tag{10}$$

the (rescaled) utility function can be expressed as

$$U_{ijs} = x'_{ijs} \beta_j + e_{ijs} \tag{11}$$

and the non-stochastic component of utility is defined conventionally as

$$V_{ijs} = x'_{ijs} \beta_j \tag{12}$$

and the posterior densities for the parameters $\{\beta_j\}, \alpha, \Omega, \{\tau_c\}$, and ϕ are obtained by observing the probability of i being chosen in the circumstance js is the standard logit probability

$$p_{ijs} = \frac{e^{V_{ijs}}}{\left(\sum_i e^{V_{ijs}}\right)} \tag{13}$$

If the observed choices are defined by $y_{ijs} = 1$ where the *ith* option is chosen in circumstance js and $y_{ijs} = 0$ otherwise, then the likelihood of all the observed choices (Y) is

$$f(Y | \{\tau_c\}, \phi, \alpha, \Omega) = \prod_i \prod_j \prod_s p_{ijs}^{y_{ijs}} \tag{14}$$

Conditionally on Λ_j and σ_{js}, the steps for generating latent variables $\{\beta_j\}$ along with α and Ω can be estimated using Markov Chain Monte Carlo (MCMC) steps as in the standard Mixed Logit (e.g., Train and Sonnier, 2005). That is, having normalised the attributes $\left(x'_{ijs} = \sigma_{js}^{-1} \dot{x}'_{ijs} \Lambda_j\right)$ the conditional distributions for β_j along with α and Ω are defined in the usual way (in terms of x_{ijs}). However, since $\{\tau_c\}$ and ϕ are estimated, the normalised attributes need to be updated at each iteration, and the posterior distributions for $\{\tau_c\}$ and ϕ are also required.

The precise priors that we use are a mean of zero for α and a diagonal covariance matrix for α with a variance of 9 for each of the elements. The precision matrix has a Wishart prior $W(I, k+4)$ where k is the dimension of the covariance matrix. The prior variance for α was set so as to be relatively uninformative for the estimates, but small enough so that the penalty for additional parameters in the model would not be overly restrictive. The posterior distributions for $\{\tau_c\}$ and ϕ therefore, conform to the following

$$
\begin{aligned}
f(\{\tau_c\}|Y,\phi,\alpha,\Omega) &\propto f(Y|\{\tau_c\},\phi,\alpha,\Omega)\, f(\{\tau_c\}) \\
f(\phi|Y,\{\tau_c\},\alpha,\Omega) &\propto f(Y|\{\tau_c\},\phi,\alpha,\Omega)\, f(\phi)
\end{aligned} \tag{15}
$$

where $f\{\tau_c\}$ and $f(\phi)$ are the prior distributions. Herein, we specify $f(\phi)$ to be standard normal and $f(\{\tau_c\}) = \prod_{c=1}^{C} I(\tau_c \in [0,1])$ where $I(.)$ denotes an indicator function which is one where the internal condition is obeyed and zero otherwise.

Estimation proceeds by iterating through the sequence of conditional draws:
i) $\{\beta_j\}|\alpha,\Omega,\{\tau_c\},\phi,Y$;
ii) $\alpha|\{\beta_j\},\Omega,Y$;
iii) $\Omega|\{\beta_j\},\alpha,Y$;
iv) $\{\tau_c\}|\alpha,\Omega,,\{\beta_j\},\phi,Y$; and
v) $\phi|\alpha,\Omega,\{\beta_j\},\{\tau_c\},Y$.

The conditional posterior distributions for the first three components (i.e., i, ii, iii) are the same as in Train and Sonnier (2005). The conditional posterior distributions for ϕ and $\{\tau_c\}$ are obtained from (15). These can be sampled using Metropolis Hastings steps with a random walk proposal density.

Finally, the framework above can be extended so as to allow for transformations of the latent normal vector $t(\beta_j)$ which is a monotonic transformation of the $(k \times 1)$ vector β_j into another $(k \times 1)$ vector. For example, we could use $t(\beta_j) = \exp(\beta_j)$. The utility function then becomes

$$U_{ijs} = x'_{ijs} t(\beta_j) + e_{ijs} \tag{16}$$

in which case $t(\beta_j)$ would be log-normal. This implies that

$$U_{ijs} = \sigma_{js}^{-1} \dot{x}'_{ijs} \Lambda_j t(\beta_j) + e_{ijs} \tag{17}$$

meaning that the marginal utilities are now $\dot{\beta}_j = \Lambda_j t(\beta_j)$. Such transformations are commonly used for price (providing a negative price is used as an attribute) or in other circumstances where there is a strong prior belief that the attribute in question yields positive marginal utility. This type of transformation makes no substantive difference to the estimation procedure as outlined above.

5. The TLS Case Study

5.1. Experimental Design

The design of a CE requires decisions to be made about i) the attributes that are included; ii) the levels of these attributes; iii) the number of alternatives in any given choice set; iv) the number of choice sets; v) whether a status quo option is included; and, vi) whether opt-outs or don't knows options are included. Then, in conjunction with assumptions about the nature of utility functions the actual choice sets presented to individuals can be generated according to efficiency criteria such as 'd-optimality' and 'balance' (Scarpa and Rose, 2008).

The decisions about i) to vi) are partly based on whether the scenarios are plausible, understandable, soluble and do not induce excessive fatigue in respondents. The majority of CEs, therefore, limit the number of attributes to be less than eight, and typically employ four or five. The number of levels for each of the attributes is generally five or less, and the levels set so as to be realistic but with good coverage of the plausible range of values. Finally, the number of choice sets given to each respondent are commonly in the range of between four and twelve.

The design used here was based on that of a previous study that has already been tested and worked well (i.e., Balcombe et al., 2010). That design was not efficient according to a criteria such as d-optimality. However, comparing this design to a d-optimal one (under zero priors for the marginal utilities) we find that there is a slight loss in d-efficiency (around 7%). But, since optimality criteria (including more recent designs for heterogeneous models) do not incorporate or allow for a number of phenomena including heterogeneous nonattendance type behaviour that we are investigating here, we did not see this as a compelling argument for replacing a tested instrument with a new one.

Our design is quite typical in that there are five attributes with three alternatives included on each choice card. Of these three alternatives one is a status quo that appears on every choice card and it does not vary. We generated a set of 24 cards which we blocked into two sets of 12. Each participant was presented with one or other of the 12 sets of choice cards.

The alternatives took the form of a food shopping basket, each with a TLS label plus the Price for the basket of goods. The TLS label is composed of a measure of Salt, Sugar, Fat and Saturates. Each nutrient took one of three levels, Green, Amber or Red, where Green is low and Red is high in terms of nutrient content. Each colour is based upon a specific quantity of the nutrient per 100 grams of food.

Based on the mix of goods in our status quo basket (which appears on every choice card) and by referring to the National Statistics (2007) publication, Family Food in 2005-06, we established the expected price of this basket of goods for an average UK household. This yielded a value of £20, and along with associated nutrient levels. The Price attribute took one of five levels. We knew from previous work that the range of alternative prices (£15 through to £30) was a sufficient dis/incentive to purchase alternative baskets, yet not so extreme as to deter all consumers from purchasing higher priced baskets.

Since the investigation of responses to colours was a component of our study we chose to have all colours within each attribute balanced across all choice cards in the two non-status quo alternatives. Since the status quo option contained three Ambers and a Red, this meant that there were a greater number of Ambers and Reds occurring overall. However, we found that requiring non-status quo alternatives to have a lower number of Ambers and Reds so that there was an overall even frequency of all colours tended to lead to some choice sets that would be highly unlikely to be chosen.

An example choice card is presented in Figure 1.

{Approximate Position of Figure 1}

5.2. Implementation of Eye-tracking

40 participants took part in the study: 28 females and 12 males aged between 18 and 25. All had normal, or corrected to normal, vision. The choice cards were presented on a colour monitor. All stimuli were presented on a white background. Eye movements were recorded using a head–mounted, video–based, eye–tracker with a recording monocularly in front of the observers' right eye. Head movements were constrained with a chin–rest, which held the participant so that their eyes were in line with the horizontal meridian of the screen. The choice made from each card were recorded through a response gamepad. The eye-tracker was calibrated at the beginning of the experiment. In order to ensure that accuracy was maintained throughout the experiment a drift correction was carried out between each card viewing. This procedure minimized the effects of slight head movements impacting on the accuracy of the eye-tracking.

Once participants were comfortable in the eye-tracker and their eye movements calibrated, they were presented with the choice cards. Participants viewed the cards for as long as they wished while we tracked their eye movements. They responded with a button press indicating which basket they would be willing to purchase. A drift correct stimulus was then shown until a button press from the participant indicated they were looking at it. The next choice card was then shown.

5.3. Alternative Specifications of the Empirical Model

The specific model that we estimated was

$$
\begin{aligned}
U_{ijs} &= V_{ijs} + e_{ijs} && (18)\\
V_{ijs} &= \left(ASC_{ijs} - \lambda_{jp} t\left(\beta_{jp}\right) p_{ijs} - \sum_k \lambda_{jk} t\left(\beta_{jk}^{red}\right) red_{ijsk} + \sum_k \lambda_{jk} t\left(\beta_{jk}^{green}\right) green_{ijsk} \right) \\
&\quad \times \exp\left(\phi'(\omega_s, z_{js})\right) \\
k &= salt,\ sugar,\ fat,\ saturates
\end{aligned}
$$

We have included ASC_{ijs} to capture any status quo effect in the response made. p_{ijs} is the Price in circumstance ijs; $red_{ijsk} = 1$ if the kth nutrient presented in circumstance ijs was Red and zero otherwise; $green_{ijsk} = 1$ if the kth nutrient presented in circumstance ijs was Green and zero otherwise; and, z_{js} is a vector defined by (4). The final term is for ϕ ($v = 1, 2, 3$, and 4) with z_{js} associated with $v = 4$, and t_s for the other parameters. The transformation $t\left(.\right)$ takes one of two forms, $t\left(\beta\right) = \beta$ or $t\left(\beta\right) = \exp\left(\beta\right)$ for all the attributes (except the status quo).

This parameterisation yields marginal utilities of Green or Red attributes relative to Amber. Attributes that are estimated to have higher marginal utility are potentially given higher visual attention. Consequently, we examine whether visual attention to attributes is positively associated with the marginal utility of attributes. If differential attention paid to attributes or colours does not translate into significantly different mean marginal utility (or vice versa), then it would not support the contention that marginal utility and visual attention are associated.

To take account of different possibilities for the estimates of marginal utility when estimating (18) we consider the following four variants (using $\alpha_k^{\mathrm{col}\,our}$ to denote the mean of β_{jk}^{colour}):

R0: Unrestricted $\alpha_k' s$;

R1: Equal Attributes $\alpha_k^{red} = \alpha_{k^*}^{red}$ and $\alpha_k^{green} = \alpha_{k^*}^{green}$ for all $k^* = salt$, $sugar$, fat, $saturates$;

R2: Colour Symmetry $\alpha_k^{red} = \alpha_k^{green}$ for all $k = salt$, $sugar$, fat, $saturates$; and, *R3*: Colour Symmetry and Equal Attributes $R1$ and $R2$.

The four variants have *R0* as the base case as it is the unrestricted model. *R1* imposes the restrictions that Reds and Greens have, respectively, the same value across nutrients (i.e., only the level of nutrients represented by colours and not nutrients themselves matter). *R2* imposes a symmetry in the mean of the latents $\{\beta_j\}$ across colours, within each of the nutrients, and *R3* imposes both these conditions.

To investigate nonattendance we used six specifications defined using the two nonattendance criteria. These are defined by equation (9) where $\delta_{1jk} = 1$ means that the *jth* individual is classified as a nonattender for attribute *k*, by stating that they did not consider the *kth* attribute, and $\delta_{2jk} = 1$ indicates that they are a nonattender according to the eye-tracking criteria (discussed below). We investigate the following six specifications for ANA:

S0: No ANA ($\tau_1 = \tau_2 = 1$);

S1: Stated ANA Only ($\tau_2 = 1$);

S2: Stated ANA = Choice ANA ($\tau_1 = 0, \tau_2 = 1$);

S3: Visual ANA Only ($\tau_1 = 1$);

S4: Visual ANA = Choice ANA ($\tau_1 = 1$, $\tau_2 = 0$); and,

S5: Visual and Stated ANA (τ_1 and $\tau_2 \in (0, 1)$).

The six ANA specifications (*S0-S5*) are compatible with each of the four restrictions on the means (*R0-R3*).

Finally, to examine scale heteroscedasticity we consider four treatments which are special cases of (4) where:

T0: No scale heteroscedasticity ($\phi_v = 0$, $v = 1, 2, 3, 4$);

T1: Sequence scale heteroscedasticity only ($\phi_4 = 0$);

T2: Visual Attention scale heteroscedasticity only ($\phi_1 = \phi_2 = \phi_3 = 0$); and,

T3: Visual and Sequence scale heteroscedasticity (no restrictions on $\{\phi_v\}$).

This means that a set of $4 \times 6 \times 4 = 96$ models were estimated for each random parameter distributional specification (i.e., normal and log-normal). The support for each restriction was evaluated by calculating the marginal likelihood ($f(Y|\mathcal{M}_m)$ as outlined in Balcombe et al. (2011) for each model $\mathcal{M}_m$. Given the large number of models being estimated, Bayesian model averaging was used to assess the support for the restrictions over all of the models.

In terms of model selection, if the model space $\{\mathcal{M}_m\}_m$, has a property (or restriction) r that defines a subset of models $\mathbf{R} = \{\mathcal{M}_m : \mathcal{M}_m \text{ has property } r\}$, and if all models within $\mathbf{R}$ are considered equally likely then $f(\mathcal{M}_m|\mathcal{M}_m \in \mathbf{R}) = 1/n_R$ (where n_R denotes the number of elements of the set $\mathbf{R}$). The marginal likelihood for $\mathbf{R}$ is

$$f(Y|\mathcal{M}_m \in \mathbf{R}) = \sum_m f(Y, \mathcal{M}_m|\mathcal{M}_m \in \mathbf{R}) = \sum_{m \in \mathbf{R}} \frac{f(Y|\mathcal{M}_m)}{n_R} \tag{19}$$

If the prior probabilities placed on two sets of restrictions are equal then the posterior odds will be the ratio of the marginal likelihoods.

6. Analysis and Results

Our results are composed of two parts. We first consider various descriptive statistics from the eye-tracking data to provide an understanding of visual behaviour. We then employ the eye-tracking data within the econometric specification we have developed.

6.1. Experimental Descriptive Statistics

6.1.1. Visual Fixations

How the survey participants have behaved in relation to fixations is summarised in Tables 1 and 2. In Table 1 the data are with respect to respondents over all 12 choice cards. The statistics show that the highest mean fixation is for Sugar and the lowest is for Price. The minimum fixations imply that for three of the attributes (Salt, Saturates and Price) there are several choice cards for which these attributes are ignored visually by some respondents (i.e. counts less than 12). Conversely, the maximum values show that some respondents have fixated on the attributes a very large number of times. This reveals a considerable degree of heterogeneity in behaviour whilst participating in the experiment. However, in terms of visual attention (measured by the number of visual fixations), respondents did not appear to radically alter their attention towards attributes in choices 1 through 12 in a systematic way.

{Approximate Position of Tables 1 and 2}

Turning to Table 2 this reports the number of times a colour occurred over the 24 choice cards and relative percentage of eye fixations. The CE was designed so that the colours in the non-status quo options were approximately balanced. The status quo option has three Ambers and a Red, meaning that the total number of occurrences within the cards is not equal for the colours. The numbers of occurrences of each colour in the experiment and the equivalent percentage are shown. The associated number of eye fixations on each of the colours, both in total and as a percentage are presented in the next two columns. These data suggest respondents had a small but significant tendency to look at Amber less frequently than it occurred. Both Green and Red were given proportionally greater attention relative to the frequency with which they occurred. We tested the hypothesis that respondents propensity to look at Green over Red or Red over Green was proportionate to the rate of their relative rates occurrence. This hypothesis could not be rejected at very high levels of significance (p=0.92). Therefore, there is little evidence that respondents were attracted to Red more than Green or vice versa, though both seemed more visually attractive than Amber.

Table 3 summarises the number of cards not fully attended. As can be seen, the number in the top left hand corner in the salt row (26), indicates that Salt was not visually attended at least once (out of the 12 choice tasks) on 26 occasions. The number to its immediate right indicates that on 18 occasions Salt was not visually attended in two or more out of the 12 choice tasks. The very bottom row shows the total number of individuals that failed to visually attend all attributes at least once, twice, and so on. What this tells us is that 35 individuals (out of 40) failed to visually attend one or more attributes at least once out of the 12 cards.

{Approximate Position of Table 3}

Table 3 demonstrates that the occurrence of choice specific visual ANA by respondents is not uncommon. For Salt, Fat and Saturates, 26, 27 and 28 out of 40 individuals respectively, failed to attend those attributes at least once within 12 choice tasks. Only five out of the 40 individuals visually attended all attributes in all experiments. At the other end of the spectrum, it is rare for individuals to repeatedly visually not attend attributes throughout the whole CE. Price was the only attribute not attended in all twelve choice cards by two individuals.

6.1.2. Relationship between Stated ANA and Visual ANA

The relationship between stated and visual attendance was explored by regressing the total number of eye fixations on each of the attributes for each respondent against stated ANA (1 = Did not attend; and 0 Otherwise). The relationship between the share of eye fixations on each of the attributes and stated ANA was also examined. For the total number of eye fixations, there were no statistically significant relationships with stated ANA. Regressions of proportion of eye fixations across attributes were slightly more significant, and are summarised in Table 4.

{Approximate Position of Table 4}

We might have expected that stated attenders would have higher average eye fixation than stated nonattenders of that attribute. However, in only two of the five attributes are these significant, and the R^2 (the share of eye attendance regressed on a stated attendance dummy) suggests an extremely poor fit for each of the five attributes. A seemingly unrelated regression, taking account of correlations in the errors by individuals, did not dramatically increase the significance. Overall, the results suggest that visual and stated ANA are very poor predictors of each other.

These findings are further supported by inspecting individual respondent data. The data show that nearly all respondents have at least one eye fix on every attribute at least once during the CE. There is only one instance of an individual who did not have a single eye fix on Price throughout the CE. Interestingly, that same individual stated that Price was one of the two attributes s/he attended. Conversely, there are numerous examples of respondents that have a higher than average share of attention on an attribute, but then state that they did not attend that attribute. For example, four respondents spent a greater than average share of their time looking at Sugar, but have stated ANA for Sugar. Overall, serial visual attendance of all attributes is not the norm. However, by the same token serial visual nonattendance is not the norm either.

6.1.3. Summary

According to a range of indicators, visual ANA is a phenomena with only a weak association with stated ANA. It may seem odd that some respondents declared that they attended attributes, when in fact they seemed to pay very little attention visually to these attributes. However, choice or information ANA does not imply that the attribute is of little importance to the irrational or semi rational respondents. Indeed, individuals can try and infer the levels of one attribute from others, without paying specific attention (i.e. looking at) to the information provided about all attributes. This can happen, for example, when attributes (e.g., Price) may be used to infer the levels of other attributes included in the CE.

6.2. Results from the Mixed Logit

6.2.1. A Definition of Visual ANA

When modelling respondents within the Mixed Logit, we require a definition of visual ANA. Total eye fixation counts are not an accurate way to assess ANA. Given that each choice set had three alternatives, and one was a status quo that did not vary across experiments, if in a given choice set, an individual did not have at least two eye fixations on a given attribute then they cannot have attended the attribute within that choice task (according to our definitions in Section 2). Thus, we have assumed that if an individual has not fully attended a given attribute in the majority of the choice tasks (e.g. over 6) then they were classified as a visual nonattender. As can be seen from Table 3, this procedure classified only six individuals as visual nonattenders. What is also interesting is that five of these individuals did not visually attend two attributes. Only one was a visual nonattender of one attribute (Price).

It is also worth remembering that we include the log of fixation duration for each individual in the scale heteroscedasticity part of the model specification. This means we can assess if individuals who spent longer considering a choice card had a lower variance attached to the Gumbel error, and, in a sense, are more certain about their choice.

6.2.2. Model Comparisons

In this section we present and discuss the results from the Mixed Logit introduced in Sections 4 and 5. We first examine the logged marginal likelihood (LMLIK) values for the specifications discussed in Section 5. We estimated all models for both normal and log-normal random parameter specifications for the four different mean restricted cases *(R0-R3)*, the six different treatments of the nonattendance data *(S0-S5)* and the four different treatments for scale variance *(T0-T3)*. Thus, in total we estimated 192 models.

Our first set of model results are presented in Table 5. These results summarise our model comparison exercise.

{Approximate Position of Table 5}

The main result reported in Table 5 is that the preferred model, specification is *R1:T1:S5*, assuming a normal distribution for the random parameters. We arrived at this result as follows.

When we compare models *R0-R3* it can be seen that the highest LMLIK is for model specification *R1*. There is little support for models *R2* and *R3* which suggests that respondents did not, on average, have the same absolute marginal utilities for Red and Green attributes (relative to Amber). This results suggests that many of the respondents did not greatly differentiate between attributes. It also suggests that the visual attention paid to specific attributes does not represent an accurate guide to how valuable an attribute is to the respondent, since visual attention was distributed equally across attributes.

Next we considered models assuming different scale heteroscedasticity *(T0-T3)*. *T1* is preferred which suggests there is little evidence that respondents who pay more visual attention to each choice task have higher or lower scale variances. These results also indicate that respondents are subject to fatigue and/or learning throughout the CE, which has previously been reported by Waldman and Savage (2008).

Finally, we consider the ANA specifications *(S0-S5)*. As we can see from Table 5 *S5* is the preferred specification. Given the preceding results this is unsurprising, since we have already seen that the two measures of ANA (i.e., stated and visual) are largely unrelated, but both improved model performance individually. Therefore, perhaps the most interesting finding is that these measures together provide largely non-overlapping but useful sources of information about respondents utility functions. Interestingly, these findings do not support Balcombe et al. (2011) who found in three out of four data sets that imposing zero utility on respondents with stated ANA improved model performance. However, the approach used in Balcombe et al. (2011) did not employ the same approach to integrating stated ANA information as that outlined in Section 4.

6.2.3. Model Estimates

We only present results for our preferred model *(R1:T1:S5)* in Table 6. To reiterate, this model has the same α across the non-price attributes *(R1)*, employs both visual and stated ANA *(S5)* and has scale sequence heteroscedasticity *(T1)*.

{Approximate Position of Table 6}

In the top part of Table 6, the mean (α) and variances $\{\Omega_{kk}\}$ of the distribution for the latent parameters are presented. In presenting the results we have changed the signs from that in equation (18) so as to reflect the direction in which each of the attributes acts on utility. These are the parameter estimates before we take account of visual and stated ANA. For the mean values in the first numerical column we can see that increased Price has a negative mean impact (-2.565) as we would expect. Amber to Red has a negative mean impact (-2.279) that is relatively larger, in absolute terms, compared to the mean impact for Amber to Green (1.322).

Next consider the bottom part of Table 6, where we report the estimates of the mean marginal utilities for each of the attributes $E\left(\lambda_{jk}\beta_{jk}\right)$. The reason these differ from the means of the latents in the top part of Table 6 is because of nonattendance. The marginal utilities are scaled by the nonattendance coefficients, along with the probabilities that an attribute was not attended (stated and visually). For this reason Fat (which has over 90% attendance visually and stated) has only a slightly lower marginal utility than the mean of its associated latent variable. The other coefficients are considerably smaller than the mean of the latent variables.

Next we consider the estimates of the ANA coefficients $\{\tau_i\}$ which are given in the second section of Table 6. The estimate of $\tau_1 = 0.653$ indicates that stated nonattenders have about 66% of the marginal utility of attenders. The coefficient for visual ANA is smaller $\tau_2 = 0.384$. In both cases the posterior distributions for these coefficients have a mass away from zero, as is reflected by the fact that the standard deviations are less than half the level of the estimates.

The scale variance parameters $\{\phi_v\}$ which capture scale heteroscedasticity are next in Table 6. The first parameter ϕ_1, being positive, indicates that overall the scale variance has fallen over the choice sequence. However, its standard deviation is also relatively large. The two sinusoidal terms are also both positive signalling a decrease in the variance in the middle phase of the CE, but with a reversal towards the end. The overall behaviour of the scale variance is best summarised graphically in Figure 2.

{Approximate Position of Figure 2}

In Figure 2 there are four lines presented. The three solid lines give the mean scale variance bounded by a 90% credible interval. As can be seen from this line, the scale variance is very high at the beginning but rapidly falls reaching its minimum at around the 4th or 5th choice card. After that, there is a gradual increase in the variance. This is consistent with the 'learning and fatigue' behaviour of respondents.

As we have already seen, visual attention did not have any significant impact on the scale variance. Additionally, we also plot the scaled average (normalised to a mean of one) eye fix duration as the dotted line in Figure 2. This shows there is little evidence of a systematic rise or fall in average eye fixation duration over the choice tasks. This illustrates that at the aggregate level, there does not seem to be a correlation between visual attention and scale variance.

Finally, our definition of visual ANA has been examined to see if it significantly influences our results. The definition of visual ANA we have employed so far is a 50% rule (i.e., if somebody did not look at the attribute in at least 6 out of 12 choice cards they were classified as a nonattender). To assess it importance we re-examined our top performing model and redefined visual

ANA as: a) respondents do not look in 25% of cases (e.g., they ignore in at least 3 out of 12 choice sets) or; b) respondents do not look in 75% of cases (e.g., they ignore in at least 9 out of 12 cases). Under both a) and b) the top model *(R1:S1:T5)* still outperforms all other models in the *R1:T1* class (with LMLIKs of -310.630 and -311.063 respectively) while the 50% rule outperforms both a) and b). Thus, we contend that or central result that both eye-tracking and stated ANA are complementary is maintained even if we modify our specific definition of nonattendance.

7.Conclusions

In this study we found that most respondents visually attended most of the attributes most of the time. However, full visual attendance of all attributes throughout the CE is uncommon as is full visual nonattendance of any attributes. If one accepts that visually fixing on objects implies that information about that object has been processed, then eye-tracking confirmed that stated ANA does not imply that a respondent has systematically ignored the information about the levels of attributes when making their decisions.

While we found some evidence of an association between nonattendance of the stated and visual forms, this was very weak. Stated nonattendance and visual forms of ANA seem to signal quite different things. Moreover, stated ANA did not appear to indicate choice ANA given the evidence from our estimated models. In this respect our results are in accordance with the majority of previous studies suggesting respondents have lower, but non-zero, marginal utility for those attributes that they state they have not attended. Thus, the model estimates together with the eye-tracking measures suggest that respondents use the stated ANA question as an opportunity to signal that something was of 'low value', but not that it played no role in their choices.

That information and/or choice ANA are much less common than is suggested by stated ANA supports the contention that a RUM approach to respondent decision making within CEs is reasonable, even if it approximates rather than accurately reflects all respondents behaviour. It is perhaps no surprise that a small number individuals appear to behave in a way that is hard to reconcile with a RUM (e.g. systematically ignoring Price). Whether such individuals should be eliminated or treated separately from the rest of the sample remains an open question. Regardless, our results show that stated and visual ANA information provide useful insights into respondent behaviour. While it would be a mistake to assume that stated ANA necessarily implies choice ANA, by using stated ANA data in the way we have, it appears that model performance can be enhanced by incorporating ANA measures.

In terms of visual attention, respondents varied widely in the visual attention they paid to attributes, and on average some attributes were paid much greater attention than others. We found little evidence that the fixation duration on colours or attributes indicated how important those attributes were. Although respondents seemed to value a move from Red to Amber more highly than the move from Amber to Green, this was not reflected in higher visual attention on Reds relative to Greens. Likewise, there was differential attention paid to attributes, but in terms of average marginal utility there was no significant difference in the means of the marginal utilities.

We did not find any compelling evidence that visual attention (in terms of fixations) had an association with scale variance. Although our scale variance appeared to have sequential heteroscedasticity, this was not reflected by higher or lower average visual attention. Moreover, by conditioning the variance on fixations we did not improve model performance. Therefore, providing the respondent has attended an attribute, looking longer or more often at an attribute does not mean it is of 'higher value'. Likewise, a respondent that pays greater visual attention overall is no more or less certain about their choices than a respondent that pays far less visual attention.

Overall the results in this paper suggest that eye-tracking is a method that promises to enhance our understanding of the cognitive processes of respondents within a CE, and to improve the estimates from models using experimental data. The evidence also suggests that further research using eye-tracking within CEs is needed. The current study was restricted to student respondents. However, eye-tracking technology is portable and the same form of study can in principle be implemented in any face to face CE. A wider demographic range of respondents may reveal different behaviours than revealed here. This study had only one stated attendance question after the completion of all the choice tasks. A question after the completion of each choice task may be more revealing as it can be matched with eye fixations by choice card. However, we are concerned that repeated stated attendance questions may induce nonattendance of attributes.

More generally, eye-tracking may be used in the visual design of CEs. Currently, little is known about the practical implications of using CE instruments that are formally the same, but different in appearance. For example, colour, size, illustrations, relative positioning of attributes, and orientation may have an effect on respondent choices. The impact of complexity of CE designs has so far been investigated using implied or stated nonattendance. Investigating complexity using eye-tracking promises to shed new light on this and other issues.

References

Aribarg, A., Pieters, R. and Wedel, M. (2010). Raising the BAR: Bias Adjustment of Recognition Tests in Advertising, **Journal of Marketing Research**, XLVII: 387-400.

Awh, E., Armstrong, K. M., and Moore, T. (2006). Visual and oculomotor selection: Links, causes and implications for spatial attention. **Trends in Cognitive Sciences**, 10, 124-130.

Balcombe, K.G., Fraser, I.M. and Di Faclo, S. (2010). Traffic Lights and Food Choice: A Choice Experiment Examining the Relationship Between Food Labels and Price. **Food Policy** 35: 211-220.

Balcombe, K.G., Burton, M. and Rigby, D. (2011). Skew and Attribute NonAttendance within the Bayesian Mixed Logit Model, **Journal of Environmental Economics and Management,** 62(3): 446-461.

Caplin, A. and Dean, M. (2008). Economic Insights from 'Neuroeconomic' Data, **American Economic Review**, 98(2): 169-174.

Caplin, A., Dean, M. and Martin, D. (2011). Search and Satisficing, **American Economic Review**, [forthcoming].

Chorus, C.G., Arentze, T.A. and Timmermans, H.J.P. (2008) A random regret minimization model of travel choice. **Transportation Research Part B**, 42(1): 1-18.

DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field, **Journal of Economic Literature**, 47(2): 315-372.

Deubel, H., and Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. **Vision Research**, 36: 1827-1837.

Fiebig, D.G., Keane, M.P., Louviere, J. and Wasi, N. (2010). The Generalised Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity, **Marketing Science**, 29(3): 393-421.

Findlay, J. M. (2009). Saccadic eye movement programming: Sensory and attentional factors. **Psychological Research**, 73: 127-135.

Franco-Watkins, A.M. and Johnson, J.G. (2011a). Applying the *Decision Moving Window* to Risky Choice: Comparison of eye-tracking and Mouse-Tracing Methods. **Judgement and Decision Making**, 6(8): 740-749.

Franco-Watkins, A.M. and Johnson, J.G. (2011b). Decision Moving Window: Using Interactive Eye Tracking To Examine Decision Processes. **Behavioral Research**, 43: 853-863.

Hensher, D.A. (2006). How Do Respondents Process Stated Choice Experiments? Attribute Consideration Under Varying Information Load, **Journal of Applied Econometrics**, 21: 861-878.

Hensher, D.A., Rose, J. and Greene, W.H. (2005). The Implications on Willingness to Pay of Respondents Ignoring Specific Attributes, **Transportation**, 32: 203-222.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics, **American Economic Review**, 93(5):1449-1475.

Knoepfle, D.T., Camerer, C.F. and Wang, J.T. (2009). Studying Learning in Games Using eye-tracking, **Journal of the European Economic Association**, 7(2-3): 388-398.

Marley, A.A.J. and Louviere, J.J. (2005). Some probabilistic models of best, worst, and best-worst choices, **Journal of Mathematical Psychology**, 49:464-480.

Marley, A.A.J., Flynn, T.N. and Louviere, J.J. (2008). Probabilistic Models of Set-Dependent and Attribute-Level Best-Worst Choice. **Journal of Mathematical Psychology**, 52:281-296.

E. Matin., E. (1974).. Saccadic suppression: a review and an analysis. **Psychological Bulletin**, 81:899–917.

McFadden, D. (2001). Economic Choices, **American Economic Review**, 91 (3): 351-378.

McFadden, D. and Train, K. (2000). Mixed MNL Models for Discrete Response, **Journal of Applied Econometrics**, 15(5): 447-470.

McSorley, E., Walker, R. and Haggard, P. (2009). The spatial and temporal shape of oculomotor inhibition. **Vision Research**, 49: 608-614.

McSorley, E. and McCloy, R. (2009). Saccadic eye movements as an index of perceptual decision-making, **Experimental Brain Research**. 198: 513 – 520.

Rayner, K. (2009). Eye movements and attention in reading, scene perception,and visual search. **The Quarterly Journal of Experimental Psychology**, 62: 1457–1506.

Reutskaja, E., Nagel, R., Camerer, C.F. and Rangel, A. (2011). Search Dynamics in Consumer Choice under Time Pressure: An eye-tracking Study, **American Economic Review**, April: 900-926.

Savage, S.J. and Waldman, D.M. (2008). Learning and Fatigue During Choice Experiments: A Comparison of Online and Mail Survey Modes, **Journal of Applied Econometrics**, 23: 351-371.

Scarpa, R. and Rose, J.M. (2008). Design Efficiency for Non-Market Valuation with Choice Modelling: How to Measure it, what to Report and Why. **Australian Journal of Agricultural and Resource Economics** 52: 253-282.

Scarpa, R., Thiene, M. and Hensher, D.A. (2010). Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services: Does it Matter? **Land Economics**, 86(4): 817-839.

Schulte-Mecklenbeck, M., Murphy, R.O. and Hutzler, F. (2011). Flashlight-Recording Information Acquisition Online. **Computers in Human Behavior**, 27: 1771-1782.

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. **Acta Psychologica,** 135: 77-99

Train, K. and G. Sonnier (2005). Mixed Logit with Bounded Distributions of Partworths, in **Applications of Simulation Methods in Environmental Resource Economics**, R. Scarpa and A. Alberini (Eds.), Kluwer Academic Publishing.

Tversky, A. (1972). Elimination by Aspects: A theory of choice. **Psychological Review**, 79: 281-299

Walker, R. and McSorley, E. (2006). The parallel programming of voluntary and reflexive saccades. **Vision Research**, 26: 2082-2093.

Wang, J.T., Spezio, M. and Camerer, C.F. (2010). Pinnocchio's Pupil: eye-tracking and Pupil Dilation to Understand Truth-Telling and Deception in Sender-Receiver Games, **American Economic Review**, 100(3): 984-1007.

Wedel, M. and Pieters, R. (2007). A Review of eye-tracking Research in Marketing, **Review of Marketing Research**, 4: 123-147.

Table 1: Fixation Counts on All Attributes

	Salt	Sugar	Fat	Saturates	Price
Mean	58.3	84.8	79.6	45.8	45.6
Std Error	6.1	7.2	7.4	5.1	4.7
Median	46.5	74	64	39.5	42
Minimum	10	23	25	3	0
Maximum	217	217	219	153	152

Table 2: Colour Frequency and Fixation Proportions

	No in Cards	No in Cards %	Eye Fixes	Eye Fixes %	SE:Eye Fixes %
Red	88	30.6	3041	33.6	0.5
Amber	135	46.9	3664	40.5	0.5
Green	65	22.6	2340	25.9	0.5
Total	288		9045		

Table 3: Visual ANA Frequency By Attribute

	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$\geq$	$=$
No of Cards not Fully Attended	1	2	3	4	5	6	7	8	9	10	11	12
Salt	26	18	11	5	3	2	1	1	1	1	0	0
Sugar	13	4	1	1	1	0	0	0	0	0	0	0
Fat	11	6	5	2	0	0	0	0	0	0	0	0
Saturates	27	19	12	10	9	5	5	5	3	3	1	0
Price	28	22	15	13	10	6	5	5	4	4	4	2
Total Non-Full Attenders	35	31	24	20	15	8	6	6	5	5	4	2

Table 4: Proportion of Eye Fixations by Stated ANA Class

Stated Attenders	Salt	Sugar	Fat	Saturates	Price
Mean	0.195	0.287	0.255	0.148	0.165
Standard Error	0.015	0.012	0.009	0.010	0.012
Stated Nonattenders					
Mean	0.174	0.266	0.262	0.112	0.116
Standard Error	0.015	0.011	0.018	0.013	0.016
Difference	0.021	0.022	-0.007	0.036	0.049
P Value	0.357	0.178	0.821	0.042	0.051
R^2	0.022	0.047	0.030	0.104	0.096

Table 5: Marginal Likelihood by Model Attributes

Distribution of Marginal Utilities	Normal	Log-Normal
R: Marginal Utility Restrictions		
R0:Unrestricted α	-322.23	-322.65
R1:Equal Attributes	-310.95	-312.17
R2:Colour Symmetry	-320.99	-324.32
R3:R1 and R2	-313.80	-317.98
T: Scale Heteroscedasticity (SH)		
T0: No SH	-315.58	-316.06
T1:Sequence SH	-311.39	-312.68
T2:Visual Attention SH	-317.09	-316.66
T3:Visual and Sequence SH	-311.85	-313.17
S: Attribute Nonattendance (ANA)		
S0:No ANA	-323.10	-322.55
S1:Stated ANA Only	-314.83	-315.94
S2:Stated ANA Only = Choice ANA	-371.94	-370.46
S3:Visual ANA Only	-318.36	-317.29
S4:Visual ANA Only = Choice ANA	-318.72	-317.85
S5:Visual and Stated ANA	-310.50	-311.78
Top Model (R1,T1,S5,Normal)	-308.942	

Table 6: Parameter Estimates Preferred Model (R1:T1:S5)

	Distribution of Latent variables			
	Mean		Variance	
	Mean	Stdv	Mean	Stdv
ASC	0.05	0.280	0.340	0.251
Price	-2.565	0.396	1.289	0.783
Salt Amber to Green	1.322	0.200	0.462	0.351
Salt Amber to Red	-2.279	0.262	0.392	0.312
Sugar Amber to Green	1.322	0.200	0.326	0.230
Sugar Amber to Red	-2.279	0.262	0.305	0.211
Fat Amber to Green	1.322	0.200	0.325	0.219
Fat Amber to Red	-2.279	0.262	0.331	0.232
Saturates Amber to Green	1.322	0.200	0.305	0.211
Saturates Amber to Red	-2.279	0.262	0.892	0.671
Nonattendance $\tau_{\mathbf{1}}$	0.653	0.068		
Nonattendance $\tau_{\mathbf{2}}$	0.384	0.168		
Scale Variance $\phi_{\mathbf{1}}$	0.596	0.382		
Scale Variance $\phi_{\mathbf{2}}$	0.697	0.264		
Scale Variance $\phi_{\mathbf{3}}$	0.497	0.181		
	Distribution of Marginal Utilities			
ASC	0.05	0.583		
Price	-2.106	1.172		
Salt Amber to Green	1.12	0.637		
Salt Amber to Red	-1.927	0.684		
Sugar Amber to Green	1.103	0.535		
Sugar Amber to Red	-1.898	0.612		
Fat Amber to Green	1.285	0.567		
Fat Amber to Red	-2.222	0.599		
Saturates Amber to Green	1.099	0.572		
Saturates Amber to Red	-1.890	0.980		

CHOICE CARD 1			
Food Basket	***Option 1***	***Option 2***	***Option 3***
Salt	Amber	Red	Green
Sugar	Amber	Green	Amber
Fat	Red	Amber	Red
Saturates	Amber	Amber	Red
Price of basket	£20	£25	£30
Click ONE and only one box			

Figure 1: Representative Choice Card

Figure 1: **Figure 2: Scale Variance**