

Using Eye-Tracking to Account for Attribute Non-Attendance in Choice Experiments

Van Loo, E., Nayga, R.M., Verbeke, W., Campbell, D. and Seo, H.S.

Accepted in European Review of Agricultural Economics, Nov 2017

Abstract

This study uses eye-tracking measures to account for attribute non-attendance (ANA) in choice experiments. Using the case of sustainability labelling on coffee, we demonstrate various approaches to account for ANA based on number of fixation count cut-offs, definitions for detecting ignored attributes, and methods for modelling ANA. Some of the sustainability attributes identified through eye-tracking measures as being 'visually ignored' were truly ignored while in none of the tested approaches price was truly ignored. The adequacy of eye-tracking as a visual ANA measure might thus depend on the attribute. Further, there are inconsistencies in identifying non-attenders using visual ANA and the coefficient of variation. Based on our results, we cannot conclude that eye-tracking always adequately identifies ANA. However, we identified several major challenges that can assist in further optimizing the use of eye-tracking in the context of ANA.

Key words: Decision-making; Eye-tracking; Sustainability labelling; Visual attention; Visual attribute non-attendance

1 Introduction

A growing body of literature applies choice experiments (CEs) as a valuation method. In a CE, respondents are asked to select their preferred alternative from a choice set¹ in which each alternative is described by attributes with varying levels. The analysis is based on the economic theory of consumer behaviour (Lancaster, 1966; McFadden, 1974), which assumes continuous preferences and thus unlimited substitutability between the attributes (Ryan and Bate, 2001). This implies that all the attributes presented as well as the trade-offs between attributes are considered (Hensher et al., 2005). However, several studies have questioned the assumption of compensatory behaviour because respondents may ignore some attributes in a choice task (Campbell et al., 2008; 2011; Carlsson et al., 2010; Hensher, 2006; Hensher and Greene, 2010; Hensher et al., 2005; 2012; Hole, 2011; Kragt, 2013; Lancsar and Louviere, 2006; Scarpa, et al. 2009; 2010). Respondents may not make the assumed trade-offs between all the attributes due to attribute non-attendance (ANA), resulting in a violation of the continuity axiom. This decision heuristic has gained increased attention in the CE literature (Hensher, 2014). Not accounting for ANA has been found to affect coefficient estimates and model performance (Campbell et al., 2008, 2011; Carlsson et al., 2010; Hensher and Rose, 2009; Mariel et al., 2013; Scarpa et al., 2009; 2010).

Mainly two methods have been proposed to identify ANA in CEs: 1) asking respondents which attributes they ignored (i.e., stated ANA) and 2) inferring ANA based on observed choices (i.e., inferred ANA). Respondents can be asked whether an attribute was ignored when making a decision at the end of the entire choice task sequence (i.e., serial stated ANA) (Alemu et al., 2013; Carlsson et al., 2010; Hensher et al., 2005; Kehlbacher et al., 2013; Kragt, 2013;

¹ In the CE literature, the set of alternatives that an individual must consider to arrive at his/her choice is called the choice set (Hensher et al., 2015). The choice task refers to the action itself of selecting the preferred alternative in the choice set.

Scarpa et al., 2013) or after each individual choice task (i.e., choice task stated ANA) (Caputo et al., 2017; Meyerhoff and Liebe, 2009; Puckett and Hensher, 2008; 2009; Scarpa et al., 2010). The disadvantage of stated ANA is that these measures are self-reported, which raises concerns about reliability (Hensher and Rose, 2009). For example, responses may be influenced by how the question is asked or interpreted. Respondents may not recall how they chose, or may not answer the attendance statement truthfully (Kragt, 2013; Scarpa et al., 2013), or may bias their answer in a socially desirable manner (Mørkbak et al., 2014). Additionally, serial stated ANA questions may be difficult to answer because they may have applied different attribute processing strategies for each choice task (Hess and Hensher, 2010; Puckett and Hensher, 2009). While asking these questions at the end of each choice task allows the respondents to indicate different ANA behaviour for each choice task, it also informs them about the researcher's interest in their attribute attendance, which may itself influence their information processing in later choice tasks. Another drawback is the additional financial cost in terms of survey time of repeatedly asking these supplementary questions as well as the increased risk of respondent fatigue, which can lead to more random decision making (Campbell et al., 2015). The standard approach in stated ANA studies is to restrict the coefficients in the utility function of attributes which have been stated as ignored to zero (Hensher et al., 2005). However, the stated ANA literature suggests that even though respondents stated to have ignored an attribute, it is possible that they may in fact have still attended to it and it may have (perhaps to a lesser extent) influenced their choice, leading to coefficients for these ignored attributes that are in fact significantly different from zero (Alemu et al., 2013; Campbell and Lorimer, 2009; Hess and Hensher, 2010). Thus, assuming attributes to be ignored based on stated ANA may lead to biased results.

Rather than relying on self-reported information regarding attribute attendance, a second method infers ANA behaviour using analytical models such as equality constrained latent class

(ECLC) models, which impose specific restrictions on the utility functions for each class by constraining some coefficients to zero for selected attributes in a certain class (Campbell et al., 2011; Caputo et al., 2013; Hensher and Greene, 2010; Kragt, 2013; Lagarde, 2013; Scarpa et al., 2009; 2013) Yet another method of inferring ANA is based on the coefficient of variation of individual-specific posterior means (Hess and Hensher, 2010; Scarpa et al., 2013; Mørkbak et al., 2014). More research has been called for on other methods to account for ANA (Caputo et al., 2017; Scarpa et al., 2013). In this study, we propose a third method based on visual ANA that is defined as visually ignoring information about attribute levels (Balcombe et al., 2015). The use of eye-tracking has been widely applied in the fields of marketing and psychology; however, it is relatively new in the field of economics. While some researchers, such as Scarpa et al. (2013), have suggested the use of eye-tracking to obtain information on ANA in CEs, limited studies have done so (Balcombe et al., 2015; Krucien et al., 2017; Spinks and Mortimer, 2016).

Visual attention is measured by eye-tracking equipment during the CE and eye-fixation counts are used as one of the indicators of visual attention. Based on the fixation counts for a particular attribute, we created a discrete measure of visual attendance by indicating whether a respondent visually attended an attribute or not. Similarly as in stated ANA studies, we apply the standard approach and develop models in which the coefficients in the utility function are restricted to zero for the attributes that are identified as visually ignored. Next, these models incorporating visual ANA are compared to a CE model in which full attendance is assumed. We then assess the performance of the visual attendance measure to identify the ignored attributes by testing whether the attributes identified as visually ignored influenced the choices. To do so, we estimated two coefficients for each attribute (attended and ignored) and test whether the ignored coefficient is different from zero, similar to what has been done in other stated ANA studies (Alemu et al., 2013; Balcombe et al., 2011; Campbell and Lorimer, 2009; Hess and Hensher, 2010).

The visual ANA study of Balcombe et al. (2015) made a few assumptions, such as requiring at least two fixations to consider an attribute visually ‘attended to’ and not allowing ANA behaviour to vary across choice sets by modelling serial ANA. We advance the investigation of visual ANA and extend the work of Balcombe et al. (2015) in three directions in relation to 1) the fixation count cut-off, 2) the definition of visual ANA detection in a choice task, and 3) the ANA modelling. Based on these three aspects, a total of six approaches or combinations are applied to incorporate visual ANA in choice models (see the overview in Table 3).

2 Material and methods

2.1 Sustainability labels on coffee

Sustainability aspects of food are credence attributes and are thus unobservable to consumers unless explicitly labelled. However, consumers may be overwhelmed with information in a shopping environment and may not pay attention to all food labels (Grunert, 2011). We apply our study to the case of sustainability labelling on coffee as coffee is one of the most popular sustainability-labelled food products. Many US coffee products carry sustainability labels such as Fair Trade (e.g. Fair Trade USA), Rainforest Alliance, and USDA Organic, which are all included in our study (see also Van Loo et al., 2015). Coffee producers are often certified for more than one type of label. For example, in 2013, 77% of the Fair Trade certified producer organizations reported holding at least one additional certification (52% Organic and 12% Rainforest Alliance) (Fairtrade International 2015, p 59). For coffee specifically, approximately 37% of Fair Trade coffee is also organic certified (Fairtrade International, 2015). Due to the proliferation of sustainability labels for coffee, coffee packages often carry several of these labels (Pierrot et al., 2011).

2.2 *Experimental design of the choice experiment*

Participants were recruited from a consumer profile database (N=6,500) *Research centre name (City, State) (omitted for refereeing purposes only)* which includes area residents. In total, 81 consumers who had purchased coffee in the two months preceding the study (March and April 2013) and who had no history of eye disease or eye surgery participated in the study. While this could be considered low for a choice experiment, the number of participants is relatively high for an eye-tracking study as recent eye-tracking studies have recruited fewer participants (e.g., 40 in Balcombe et al., 2015; 18 in Bialkova et al., 2013; 59 in Danner et al., 2016). Each participant was given a \$20 gift card. Approximately half (53%) of the participants were female (Table 1). Each age and income category is represented. The sample is slightly biased towards participants with higher education.

----Insert Table 1-----

All coffee products in the experiment were ground medium roast coffee, the most popular type of coffee in the US (Mintel, 2012). The coffee products were described using a combination of five attributes: four sustainability labels and price (Table 2). Each of the sustainability labels had two levels (present or not present) and the price attribute had four levels based on a store check in food stores in *City (State) (omitted for refereeing purposes only)*, USA in April 2013.

----Insert Table 2-----

Given the four attributes and their levels, a full factorial design would have resulted in 64 ($2^4 \times 4$) alternatives resulting in 4,096 choice sets - each consisting of two alternatives. The CE design followed Street and Burgess (2007). We used a fractional factorial design and the generator ([1 1 1 1 1]) to obtain the design of eight choice sets (Burgess, 2007; Street and Burgess, 2007) with an efficiency of 97.6%. To increase the similarity to a real shopping experience, a no-buy alternative was added. Hence, in each choice set, participants were presented with two types of ground medium roasted coffee as well as a no-buy alternative (Figure 1). Due to the hypothetical

nature of our CE, a cheap talk script was presented prior to the choice tasks. The location of the labels (USDA Organic, Rainforest Alliance, Fair Trade, Carbon Footprint) on the package (from left to right) was randomized for each of the eight choice sets to avoid an order effect due to label location. In addition, the randomization for each of the choice sets was repeated ten times, resulting in ten different surveys. Each respondent was randomly assigned to one of these ten surveys. Additionally, within each survey, the eight choice sets were randomly presented to avoid order effects.

----Insert Figure 1-----

2.3 Experimental procedure for the eye-tracking experiment

While addressing the eight choice sets, the participants' visual attention was recorded using a contact-free eye-tracking device (model: RED, SensoMotoric Instruments GmbH (SMI), Teltow, Germany) located in a panel beneath a 56 cm computer screen (474 mm by 297 mm and screen resolution of 1680 px by 1050 px) (see Appendix A). The approximate distance between the display monitor and each participant's head was 70 cm. The sampling rate and tracking resolution of the eye-tracking device were 120 Hz and 0.03° , respectively. Visual stimuli, which are the images presented on the screen representing the choice set, were randomly presented using stimulus presentation software (Experiment Suite 360^o™, SensoMotoric Instruments GmbH, Teltow, Germany).

Before the CE task, participants received instructions and the eye-tracking device was individually calibrated using a five-point calibration method with a low mean tracking error (less than 0.4°). After successful calibration, two warm-up choice sets were presented to familiarize each participant with the experimental procedures. As in Balcombe et al. (2015), participants knew that eye-tracking was used; however, they were not aware of its purpose. As visual stimuli, the choice sets with the two coffee packages were presented and participants chose the preferred alternative. After the two warm-up choice sets, they were randomly assigned

to one of the ten surveys, and they then answered all eight choice set questions, which were randomly presented to them. Between the choice tasks (i.e., during the inter-stimulus intervals), participants were asked to maintain their fixation on a central black cross against a white background for approximately 8 s. Similar to the procedure used in Balcombe et al. (2015), the participants viewed each choice set as long as they wanted before indicating their choice. The 81 participants each made eight choices. However, for three choice sets, the choice was not indicated, resulting in a total of 645 choices. On average, the participants spent 73 s on all eight choice tasks combined (without the inter-stimulus intervals), which equates to an average of less than 10 s per choice task. The average tracking ratio, i.e. “the number of non-zero gaze positions divided by the sampling frequency multiplied by run duration, expressed in percent” (SMI, 2016 p 335), across our sample is 89.5%.

2.4 Eye-tracking measures

Areas of interest (AOI) were defined on the coffee packages (Figure 1) corresponding to the five attributes. Using the eye-tracking software (BeGaze™, ver. 3.0, SensoMotoric Instruments GmbH, Teltow, Germany), fixation counts were calculated for the five AOIs in each of the eight choice sets. The fixation count is the number of times the participant fixated his or her gaze on the AOI. More fixations are an indication that an area is more noticeable or more important to the viewer than other areas (Poole et al., 2005). The number of fixations within the AOI has been considered a reliable measure for the visual attention given to that AOI (Bialkova and van Trijp, 2011). The low speed event detection method (suggested for <200 Hz) was selected in BeGaze for the fixation detection. In this method, the fixation is the primary event, and other events are derived from it. The method uses two specific detection parameters: a minimum fixation duration of 80 ms and maximum dispersion of 100 px. ‘The minimum fixation duration defines the minimum time window in which the gaze data is analysed. Fixations smaller than the

time window will not be caught' (SMI, 2016, p 317). The low speed event detection method uses a dispersion-based algorithm. For details on this algorithm, we refer to the BeGaze Manual 3.6 (SMI, 2016, p 317). For each stimulus, the first fixation was excluded. According to Holmqvist et al. (2011), this approach is often used because the fixation position at stimulus onset has not been influenced by the stimulus content.

2.5 Accounting for visual ANA

Fixation counts were obtained for each AOI and are a measure of visual attention. In this study, three aspects are taken into account when going from fixation counts to incorporating visual ANA in the choice modelling: 1) cut-off point for the fixation count, 2) definition of visual ANA to identify an attribute as ignored in a choice task, and 3) modelling approach for visual ANA.

2.5.1 Fixation count cut-off

Balcombe et al. (2015) assumed that at least two fixation counts are required to consider an attribute 'attended to'. However, even one fixation count can signify that the person fixated on the information and thus may have attended to the information. Hence, in addition to the arbitrary assumption by Balcombe et al. (2015) that two fixation counts are required to consider an attribute visually attended to, we also used a fixation count of one as a less strict threshold to consider an attribute as being visually attended to.

2.5.2 Defining visual ANA

To define visual ANA, we used the fixation count as a measure of visual attention and created the discrete measure 'visual ANA'². This discrete measure indicates whether the attribute is

² Visual attention to an attribute is 'a continuous measure of the degree to which a respondent evaluates the attribute', while attendance is 'a discrete measure indicating whether respondents will be considered to have attended an attribute or not' (Balcombe et al., 2015, p 449).

identified as being ignored or not in a particular choice task. We used two definitions to identify an attribute as being 'ignored' in a particular choice task.

A) Based on visual attention to a specific attribute in the choice set as a whole (which we refer to as *Definition A*)

This definition was used by Balcombe et al. (2015), who considered an attribute to be ignored in a choice task if the fixation count for the attribute summated over the alternatives within one choice set was below the cut-off. Thus, the fixation count for one attribute is calculated for the choice set as a whole.

B) Based on visual attention to a specific attribute in each of the alternatives within the choice set (which we refer to as *Definition B*)

An attribute is judged to be ignored in a given choice task if the attribute was ignored (fixation count less than the cut-off) in both alternatives (if the attribute was present in both alternatives). Rather than defining an attribute as ignored in a choice task based on the choice set as a whole (i.e., the fixation count for one attribute was summated over all alternatives within one choice set) as implemented in Balcombe et al. (2015), this definition is based on visual attention to an attribute in both the non-no-buy alternatives within a choice set. Thus, in this second definition, visual attention to the attribute in each alternative is taken into account.

Note that the use of definition A or B will only have an impact on the price attribute, as it is present in both non-no-buy alternatives, and the alternatives never share any of the quality attributes. We assumed that respondents, when looking at one package, did not infer which labels the other package had.

2.5.3 *Modelling approaches for visual ANA*

Two modelling approaches were used to account for visual ANA, one at the respondent level (serial visual ANA) and one at the choice set level (choice task visual ANA). Serial ANA refers

to classifying a respondent as an attender or non-attender for a particular attribute for the entire choice task sequence, while choice task ANA allows for differences in attendance across choice tasks. Serial and choice task visual ANA are similar to serial and choice task stated ANA but instead of stated attendance, visual attendance is used.

1) Serial visual ANA

Following Balcombe et al. (2015), we classified a respondent as having visually ignored an attribute over the whole sequence of choice tasks if the participant ignored a given attribute in more than half (i.e., more than four) of the choice tasks. Thus, using serial visual ANA, a person is either a non-attender or an attender of an attribute for the entire sequence of choice tasks in the CE.

2) Choice task visual ANA

Several authors, however, have advocated that respondents' processing strategies may change as they progress through a sequence of choice tasks, meaning that their tendency to ignore attributes may not be consistent throughout a panel of choices (Hess and Hensher, 2010; Meyerhoff and Liebe, 2009; Puckett and Hensher, 2009; Scarpa et al., 2010). Hence, it may be important to allow for varying ANA behaviour from one choice task to another. Choice task visual ANA allows visual ANA to vary across choice tasks. When a respondent visually ignored a given attribute in a choice task, this attribute is characterized as non-attended for that particular choice task. Therefore, for each choice set and for each attribute, a participant is classified as having attended or not attended to the attribute.

For each of the two visual ANA definitions to detect ignored attributes (*Definition A* and *Definition B*), both modelling approaches – serial visual ANA (S) and choice task visual ANA (CT) – were applied. This leads to four combinations – *defA-CT*, *defA-S*, *defB-CT* and *defB-S* (Table 3) – when a fixation count of two is applied as the cut-off. Definition A combined with

the serial ANA modelling method with a fixation count of two as the cut-off (*defA-S*) is the approach used by Balcombe et al. (2015).

For the cut-off fixation count of one, definitions A and B result in the exact same discrete measure of visual ANA; thus, no distinction is made between definitions A and B (see Table 3). Therefore, a fixation count of one as the cut-off results in two additional combinations or approaches (*FCI-S* and *FCI-CT*).

----Insert Table 3-----

2.6 Discrete choice models

While the multinomial model (MNL) assumes homogeneity in consumer preferences, we assume that heterogeneity may be an issue in analysing consumer preferences for food labelling (Bonnet and Simioni, 2001; Van Loo et al., 2014). Therefore, a random parameter logit (RPL) model was estimated (with 500 Halton draws) that allows for random taste variation and for the panel structure, given that each respondent made eight choices. This approach results in the estimation of a mean and standard deviation for each of the random taste parameters. For simplicity, we assume price to be a fixed coefficient, which is a widely practised specification in the field (Caputo et al., 2013; Layton and Brown, 2000; Lusk and Schroeder, 2004; Morey and Rossmann, 2003; Revelt and Train, 1998). This restriction allows the distribution of the WTP to be easily calculated from the non-price coefficients. We further assume that the coefficients of the four sustainability labels follow a normal distribution (Caputo et al., 2013; Lusk and Schroeder, 2004).

Two additional modelling issues are taken into account – the correlations across taste parameters and across utilities – to make the estimates more robust and consistent with consumer choice behaviour (Barreiro-Hurle et al., 2010; Gracia et al., 2012; Gracia, 2014). To allow for dependence across tastes, no restrictions were applied to the correlations among the random parameters. Additionally, because the design consists of two designed alternatives and

one no-buy alternative, correlations across utilities may exist (Scarpa et al., 2005). The no-buy alternative is truly experienced, while the designed alternatives can only be imagined. Therefore, the utilities of the buying alternatives are likely to be more correlated among themselves than with the no-buy alternative. To account for this correlation pattern, we employed an RPL model with an error component (RPL-EC) (Scarpa et al., 2005; 2007), whereby both designed alternatives share an extra error component that is a zero-mean normally distributed random parameter.

Specifically, with our attributes, the utility that individual i obtains from alternative j at choice situation t takes the following form:

$$U_{ijt} = \beta_0 No_Buy_{ijt} + \beta_1 Organic_{ijt} + \beta_2 Rainforest_{ijt} + \beta_3 Fairtrade_{ijt} + \beta_4 Carbonfootprint_{ijt} + \beta_5 Price_{ijt} + \eta_{ij}(1 - No_Buy_{ijt}) + \varepsilon_{ijt},$$

where j pertains to alternatives A, B and C. No_Buy_{ijt} is an indicator variable that takes the value of 1 when the no-buy alternative is chosen and 0 when either product profile A or B is selected. β_0 is an alternative-specific constant representing the no-buy alternative. Price is the price of a package of 12 ounces of coffee. η_{ij} is the zero-mean normal error term, or the error component term, which is only associated with alternatives that portray a purchase decision and is absent in the utility of the no-purchase alternative. ε_{ijt} is the unobserved random error term.

The marginal WTP values are calculated as a negative ratio, where the numerator is the estimated mean values of the coefficients associated with a particular sustainability label and the denominator is the price coefficient. The data were analysed using NLOGIT 5.0 by Econometric Software Inc. (Greene, 1990).

2.7 Accounting for ANA

The standard approach to account for stated ANA is to restrict the coefficient in the utility function to zero for the attributes that the respondents stated as ignored, which results in the removal of the attribute from the choice consideration (Hensher et al., 2005). This method has

been incorporated into the NLOGIT 5.0 software by coding an attribute as -888 if it is not attended to and assigns a zero to the attribute coefficients rather than to the attribute levels (Greene, 2012). This approach has been applied in several studies on stated ANA (Alemu et al., 2013; Hensher et al., 2005; Hensher and Rose, 2009; Kragt, 2013; Scarpa et al., 2013). We used the same approach, using a dummy variable to denote whether the attribute was visually ignored (visual ANA). For the choice task ANA, this is defined at the choice set level (choice task ANA) or at the respondent level (serial ANA). We did not collect any data on stated ANA.

For each of the six combinations (*defA-CT*, *defA-S*, *defB-CT*, *defB-S*, *FCI-S*, *FCI-CT*), we estimated a visual ANA model in which the coefficients of the visually ignored attributes are restricted to zero. In addition, a full attendance model in which all attributes are assumed to be attended to was estimated. First, MNL models were estimated (Appendix A). We then also estimated the RPL-EC models, and given that they are behaviourally more appropriate and outperform the MNL models in terms of model fit, we only discuss the results of the RPL-EC models (Table 5).

2.8 *Are the attributes identified as visually non-attended actually ignored?*

We examine whether the attributes identified as visually non-attended are in fact also truly ignored³ when respondents are making the choice by estimating the coefficient of ignored attributes (section 2.8.1). In addition, we compare the results of serial visual ANA with results from the coefficient of variation method (section 2.8.2), an inferred method to identify non-attenders.

2.8.1 *Coefficients of ignored attributes*

The coefficients of the ignored attributes are no longer restricted to zero; but are freely estimated. In the stated ANA literature, some studies have indicated that people reporting to

³ 'Truly ignored' is related to the coefficient of 'visually ignored' attributes being zero.

have ignored a certain attribute may have a marginal utility for that attribute that differs from zero. Hence, respondents who stated that they ignored an attribute may have actually considered it (Carlsson et al., 2010). As a result, instead of restricting the coefficient of ignored attributes to zero, some stated ANA studies estimate two coefficients for each attribute: one for those who stated that they attended to the attribute and one for those who stated that they did not attend to the attribute (Alemu et al., 2013; Campbell and Lorimer, 2009; Hess and Hensher, 2010; Scarpa et al., 2013). Similarly, we estimate models with two coefficients for each attribute: one for visually ignored attributes and one for visually attended attributes. If the visually non-attended attributes were truly ignored, the corresponding coefficient should not differ statistically from zero. Because of the number of coefficients to be estimated and our limited sample size, MNL models were estimated.

2.8.2 *Coefficient of variation of individual-specific coefficient distributions*

Using the serial visual ANA, a respondent is identified as an attender or a non-attender for a particular attribute. For each respondent, the allocation of having ignored an attribute or not based on serial visual ANA is compared with the allocation of this respondent as an attender or non-attender based on the inferred method using the coefficient of variation⁴.

Following Hess and Hensher (2010), we attempted to infer whether a respondent ignored a particular attribute or not (thus, inferred serial ANA) by analysing the individual-specific coefficient distributions that have been conditioned on observed choices. For additional details, we refer readers to the NLOGIT reference guide (Greene, 2012, section N29.8). Based on the RPL model in Table 5, the means and standard deviations for the conditional distributions were calculated for each coefficient of the random parameters. Rather than using the conditional mean to infer whether a respondent ignored an attribute or not, Hess and Hensher (2010) suggested

⁴ We could not apply ECLC models because the limited sample size caused issues with them.

using the coefficient of variation as an inferred method to identify ignored attributes. The coefficient of variation is the ratio between the standard deviation and the mean of the conditional distribution'. This measure is used by Hess and Hensher (2010, p 786) to incorporate uncertainty into the conditional distributions and gives an indication of whether the conditional mean is indistinguishable from zero. Hess and Hensher (2010) reported that this approach is better than using the conditional mean because a respondent may have a low sensitivity to an attribute without actually ignoring it. Hence, relying only on a low mean to allocate a respondent to the ignored group might be incorrect, and therefore using the coefficient of variation is suggested. Following Hess and Hensher (2010), we allocate respondents with a coefficient of variation of two or above for a certain attribute to the ignored group for that attribute. Subsequently, we evaluate whether the identification of respondents as having ignored or not ignored an attribute based on the serial visual ANA matches with the identification based on the coefficient of variation.

3 Results and discussion

3.1 Visual attribute non-attendance frequency

In our specifications of visual ANA (definition A or B), the modelling approach for visual ANA (serial or choice task) as well as the fixation count cut-off determine the ANA frequency. The proportions of ANA for each of the attributes and for the six different combinations are presented in Table 4.

----Insert Table 4-----

For serial visual ANA, the proportions of visual non-attenders for the sustainability labels range from 41% to 56% of the total number of participants for a fixation count of two, while the proportions range from 12% to 23% for a fixation count of one. For both fixation counts, the number of respondents who ignore the price is lower than for the sustainability labels. Using a fixation count of two as the cut-off, only 12% and 23% of the respondents were

classified as visual non-attenders for price for Definition A and Definition B, respectively⁵. For a fixation count of one, only 5% of the respondents were classified as ignoring the price.

When applying choice task visual ANA, the number of choice tasks in which the sustainability labels were ignored ranges from 52% to 60% for a fixation count of two as the cut-off and from 28% to 30% for a fixation count of one as the cut-off. Price was ignored in 25% or 41% of the choice tasks depending on the ANA definition applied for the case of a fixation count of two as the cut-off and in 10% of the choice tasks for a fixation count of one.

3.2 Standard ANA approach

Similar to the standard approach in stated ANA, models are estimated with the parameters for the visually ignored attributes being constrained to zero; i.e., four models for the combinations with a fixation count of two: serial and choice task visual ANA based on definition A (*defA-CT*, *defA-S*) and based on definition B (*defB-CT* and *defB-S*) (Table 5) and two additional models for a fixation count of one: choice task and serial visual ANA modelling approaches (*FCI-S* and *FCI-CT*). The full attendance model (*full-AA*) pertains to the estimation assuming full attribute attendance and is included as a benchmark. To allow for heterogeneous preferences among the respondents and correlation across utilities, RPL-EC models were estimated (Table 5). The MNL model estimations are included in Appendix B.

----- Insert Table 5 -----

In all of the models, the coefficient of the no-buy alternative is negative and statistically significant, indicating that participants increase their utility when choosing one of the proposed

⁵ The visual attendance towards price depends on the definition applied because price was presented in each of the two buying alternatives. For the definition based on both alternatives (definition B), price is considered to be attended to in a choice task if in both alternatives it has a fixation count of at least two as opposed to a fixation count of at least two for the prices presented in the choice set as a whole (definition A).

coffee alternatives compared with the no-buy alternative. In all of the models, the hypothesis of correlation across utilities is verified because the standard deviation of the error component (η_{ij}) for the purchase alternatives is statistically significant. Correlations across the random parameters were also allowed. The Cholesky matrices are presented in Appendix C. The coefficients of the attributes have the expected signs. The price coefficient is negative, as expected, and statistically significant at the 0.01 level, indicating that consumers' utility decreases with increasing price.

In the *full-AA* model, the coefficients of Organic, Rainforest Alliance and Fair Trade are significant, implying that respondents' utility increases when one of these labels is present on a coffee package. The results show that USDA Organic is the highest valued attribute, resulting in the strongest utility increase. The USDA Organic label is preferred over Rainforest Alliance and Fair Trade. The *full-AA* model has significant standard deviations of the random parameters (except for Rainforest Alliance), indicating the presence of considerable unobserved heterogeneity in taste preferences across the respondents.

Turning to the standard ANA approach in which the parameters of the visually ignored attributes are restricted to zero (*defA-CT*, *defA-S*, *defB-CT*, *defB-S*, *FCI-S* and *FCI-CT*), we find that most of the parameters for the considered attributes are significant at the 5% or 1% level. In all six models, the coefficient of USDA Organic is the largest. Carbon Footprint is not significant for all models except for the choice task modelling approach with a fixation count of two as the cut-off.

While the standard deviations of the random parameters for Fair Trade and USDA Organic of the full attendance model were significant at 1%, this is no longer the case when

accounting for visual ANA⁶. While the full attendance model with significant standard deviations shows preference heterogeneity, accounting for visual ANA captures an important part of the heterogeneity across participants. This result illustrates that confounding between ANA and preference heterogeneity might be an issue (Hess et al., 2013), and thus preference heterogeneity may be incorrectly interpreted when ANA is not addressed, which further illustrates its importance (Hess et al., 2013).

The full attendance model outperforms the model in which the coefficients of the ignored attributes are restricted to zero, as illustrated by the decrease in model fit for the visual ANA models (a decrease in log likelihood and an increase in the AIC and BIC statistics) (Table 5). A worse model fit could be due to treating attributes as having a zero parameter when some are not actually ignored.

3.3 Are the attributes identified as visually non-attended actually ignored?

3.3.1 Coefficients of ignored attributes

We test whether the attributes identified as visually non-attended truly have coefficients that are equal to zero by estimating them freely, which leads to model estimations with two coefficients for each attribute (considered and ignored) that are referred to as *defA-S2*, *defB-S2*, *defA-CT2*, and *defB-CT2* as well as *FCI-S2* and *FCI-CT2* for fixation counts two and one as the cut-off, respectively (Table 6). Of the 30 ignored coefficients⁷, 17 are not significantly different from zero. In the cases of the Rainforest Alliance, Fair Trade and Carbon Footprint labels, being identified as visually non-attended using one of the six approaches means that these attributes

⁶ As noted by a reviewer, this could also be due to the relatively small sample size. Only the observations in which the attribute was attended to are used for the coefficient estimation, while for the other observations, the coefficients are set to zero.

⁷ Five parameters were estimated (USDA Organic, Rainforest Alliance, Fair Trade, Carbon Footprint and Price) for each of the six approaches, resulting in a total of 30 parameter estimations.

were truly ignored, with the exception of Rainforest Alliance in the choice task ANA modelling approach with a fixation count of two.

----- Insert Table 6 -----

For all model estimations using the serial ANA modelling approach, the coefficients of the ignored attributes indicate that respondents who were identified as visual non-attenders for the Rainforest Alliance, and Fair Trade labels truly ignored these attributes⁸. For the serial ANA modelling approach with a fixation count of one (*FCI-S2*), the ignored coefficient of USDA Organic is also not significantly different from zero. Thus, for the serial ANA approach with a fixation count of one, respondents who were identified as visual non-attenders for one of the sustainability labels (USDA Organic, Fair Trade, or Rainforest Alliance) truly ignored these attributes.

For all model estimations using the choice task ANA modelling approach, the coefficients of the ignored attributes indicate that in the choice tasks in which we considered Fair Trade and Carbon Footprint (and Rainforest Alliance for a fixation count of one) as visually ignored, they were indeed truly ignored. Thus, these choice tasks were answered as if the visually ignored attribute was not present in the choice set. For these attributes, restricting the coefficient to zero if it was visually ignored was appropriate and resulted in the removal of the attribute from the choice consideration.

While 17 out of 30 ignored coefficients were not statistically different from zero, 13 out of the 30 estimated ignored parameters were significantly different from zero; thus, setting the coefficients of these parameters to zero may not be appropriate. In all six ANA models, the coefficient of ignored price was significant at the 1% level. In all ANA models except *FCI-S2*, the coefficient of ignored USDA Organic was also significant. When using a fixation count of

⁸ For Carbon Footprint, the considered coefficients are insignificant. Therefore, being wrongly classified as having ignored Carbon Footprint may result in an insignificant coefficient.

two as the cut-off, the choice task ANA model also had significant coefficients of the ignored Rainforest Alliance. Whereas assuming that visually ignored attributes were truly ignored was appropriate for all sustainability labels in the serial ANA approach with a fixation count of one and for all sustainability labels except USDA Organic in the serial ANA approach with a fixation count of two, this was not the case for price. Therefore, it is inappropriate to constrain the coefficients of the visually ignored price and also in some cases the sustainability labels USDA Organic and Rainforest Alliance to zero. This finding is important because it indicates that some attributes that were classified as visually non-attended based on the ANA definitions and modelling approaches were actually not ignored and could have influenced the choice. It is also important to acknowledge that there is a difference in the meaning of having coefficients for the ignored attributes significantly different from zero based on either stated ANA or visual ANA. When based on stated ANA, it refers to saying that an attribute was ignored while in fact it did (possibly reduced) impact one's choice (Campbell and Lorimer, 2009). For visual ANA, it refers to being identified as not having looked at the attribute while it (possibly reduced) impacts one's choice.

The statistical differences between the two coefficients (considered and ignored) were tested using the Krinsky and Robb method (Krinsky and Robb, 1986; 1990) and suggest some differences in behaviour (Table 6). For price, the coefficient of attended price was statistically lower (more negative) than the coefficient of ignored price (except for FC1-S2, where the difference was not statistically significant). The ignored coefficients of USDA Organic were significantly different from zero in five of the six models. Most of the ignored coefficients of USDA Organic were also significantly smaller than the considered coefficients of USDA Organic. For the choice task ANA models with a fixation count of two (*defA-CT2* and *defB-CT2*), the Rainforest Alliance ignored coefficients were also significantly different from zero and statistically lower than the coefficient of attended Rainforest Alliance. This result indicates

that a classification as visually ignored, on average, leads to a lower utility for the USDA Organic label (and also for the Rainforest Alliance label for two of the models) and a less negative utility for price. These less negative (price) or positive (USDA Organic, Rainforest Alliance) coefficients of the attributes identified as ignored may have two possible explanations. First, the Price, USDA Organic and Rainforest Alliance attributes that may be classified as visually non-attended may in fact not have been truly ignored when choosing the preferred alternative. Instead, respondents who paid less attention to these attributes received less negative (price) or less positive (USDA Organic) utility from these attributes.

Second, it is possible that the smaller coefficient for the ignored subset might be a combination of truly ignored attributes (zeros) and attended attributes, which therefore results in a lower value for the coefficient. This could be an artefact of the experimental design or due to a number of reasons including tracking error and/or wrongly identifying ANA due to tracking error and/or peripheral vision effects as explained in the next paragraph.

Balcombe et al. (2015) noted that people must look long enough at information for it to be processed. However, our results show that measuring whether an attribute is visually attended to or not might be attribute-dependent. For example, the number of fixations needed to visually attend to attributes may differ depending on the attribute itself. Price and USDA Organic (and in some cases Rainforest Alliance) are attributes that were attended to when we identified them as being ignored. This result may be due to the respondents' familiarity with these attributes. Therefore, even when we defined them as visually non-attended, they were not actually ignored. Wrongly assuming a coefficient to be zero also explains reduced model fit (see section 3.4). In addition to familiarity, issues specifically related to the eye-tracking technology may provide an explanation for why there are significant coefficient estimates for some of the ignored parameters. While fixating on one attribute, another attribute might be viewed simultaneously and interpreted without truly fixating on it. If the participants did not fixate on information

presented in an AOI, it does not mean that they were not aware that it was there (Bergstrom and Shall, 2014, p 6), since fixations only report visual attention taking place in the foveal vision and not in the parafoveal and peripheral vision. While people’s primary attention is focused on what they see in the foveal vision, they might still grasp information presented in other parts of their visual field. Some authors (Henderson and Hollingworth, 1999; Henderson et al., 2003) have reported that a functional field of view can be 4 degrees, while others (Holmqvist et al., 2011) have suggested using a margin of 1 to 1.5 degrees⁹. We also cannot rule out the possibility of a measurement error due to the eye-tracker as the reason why the coefficients of the ignored parameters are significant, as it would be ‘unnatural’ for people to decouple visual attention and fixations in this manner¹⁰. More information on the limitations and challenges of eye-tracking are reported in a separate section (section 5). Visual ANA is a discrete measure that indicates whether an attribute was ignored or not. However, it may be difficult to make this distinction due to 1) the use of fixations, 2) the AOIs’ sizes, 3) the definitions for visual ANA, and 4) the influence of familiarity. These aspects are discussed in detail in section 5.2.

3.3.2 *Coefficient of variation*

In this second method, we evaluate whether respondents who were identified as having ignored or not ignored an attribute based on the serial visual ANA have the same allocation (ignored versus not ignored) based on the coefficient of variation method (Table 7).

First, the rates of ignoring the different attributes between visual and inferred ANA are compared (Table 7). The proportions of respondents ignoring Fair Trade, USDA Organic,

⁹ As suggested by a reviewer, the functional field of view is likely different for illustration and text, with a smaller functional view for text (such as price) as compared to illustrations (such as the sustainability labels).

¹⁰ We thank an anonymous reviewer for pointing out this possibility.

Rainforest Alliance and Carbon Footprint using this inferred method¹¹ are 18.5%, 14.8%, 6.2%, and 35.8%, respectively. The inferred method has much lower rates of ignoring compared to the visual ANA with a fixation count of two (Table 7), except for Carbon Footprint, which is quite similar. For example, out of the 81 respondents, 15 (18.5%) were considered as having ignored the Fair Trade attribute based on the inferred method and 33 (40.7%) were considered as having ignored the attribute based on serial visual ANA with fixation count 2. The proportions obtained by the inferred method are more in line with those obtained through visual ANA with a fixation count of one as the cut-off, except for Carbon Footprint.

----- Insert Table 7 -----

In addition to the actual rate of ignoring attributes, a comparison of the allocation of the respondents identified as visually ignoring an attribute is of interest (Table 7). For example, for the attribute Fair Trade, 49 respondents are classified the same for both the inferred and the visual ANA with fixation count 2 (41 attenders and 8 non-attenders). A large proportion of the respondents who were identified as having ignored an attribute based on the serial visual ANA were not identified as ignoring this attribute by the inferred method. This shows that there is little agreement regarding the identification of serial non-attenders between the visual and inferred approach. This may suggest that a large proportion of the respondents who were identified as having ignored an attribute using the serial visual ANA did not actually ignore it.

It is clear that there are differences in the allocations into the ignored group between the serial visual ANA and inferred ANA. However, it remains uncertain which method is the most accurate in identifying ignored attributes. Hess and Hensher (2010) compared this inferred method (i.e., the coefficient of variation approach) with stated ANA and found large differences

¹¹ The inferred method cannot estimate the percentage of respondents ignoring price since price is considered fixed.

or inconsistencies between the two approaches in terms of rates of ignoring and in the allocation of respondents into the two groups (attenders versus non-attenders). The inferred method also has drawbacks such as the use of the threshold value of a coefficient of variation of two is an arbitrary threshold (Hess and Hensher, 2010). As mentioned by Hess and Hensher (2010), more research is needed into how to allocate respondents to the ignored group based on the coefficient of variation, such as defining a less arbitrary threshold for the coefficient of variation method.

3.3.3 *Model fit across estimations*

When comparing the RPL model fits, restricting the coefficients of the ignored attributes to zero results in a decrease in model fit (a decrease in log likelihood and an increase in the AIC and BIC statistics) compared with the full attendance model (Table 5). A worse model fit is likely due to treating attributes as having a zero parameter when some are not actually ignored (which was confirmed by the model estimations with two coefficients per attribute). When the ignored coefficients are freely estimated, the model fit is better than for the full attendance model. For both fixation cut-offs, the serial ANA modelling approach results in a better model fit than the choice task ANA modelling approach. For models with a fixation count of two, defining an ignored attribute based on the whole choice set (definition A) results in a better model fit than defining it based on the alternatives (definition B).

Several studies on stated ANA have reported an improvement in model fit when constraining ignored coefficients to zero (Campbell et al., 2008; Hensher et al., 2005; Kragt, 2013). However, similar to our study, Alemu et al. (2013) reported a decrease in model fit when restricting the coefficients to zero. This decrease in model fit could also be attributed to the number of observations that are essentially excluded from contributing to the likelihood function (Alemu et al., 2013), which may also explain why we see a decrease in model fit when moving from serial to choice task ANA, as the percentage of ignored attributes also increases (Table 4).

The model fits in terms of the AIC and BIC statistics across the six models that used two coefficients per attribute are shown in Table 6. First, compared with the full attendance model (Appendix B), models in which two separate coefficients were estimated have a better model fit compared with the model assuming full attendance. In addition, for each approach, the model with two coefficients (Table 6) outperforms the model in which the ignored coefficients were set to zero (Appendix B). Setting the ignored coefficients to zero when in fact they are statistically different from zero worsens the model fit.

Since not all coefficients that were identified as ignored were, in fact, ignored based on estimates of these coefficients (section 3.3.1) and based on the coefficient of variation (section 3.3.2), we can conclude that eye-tracking did not always successfully identify visual non-attenders for all attributes. This is also confirmed by the decrease in model fit for most models when constraining the ignored coefficients to zero compared with the full attendance models. We discuss the potential challenges and limitations of using eye-tracking for ANA research in section 5.

3.4 WTP

We cannot always assume that the respondents did not look at the attribute just because they did not care about it; i.e., we do not know the reason why they did not consider the attribute when making their choices. For instance, some people may have visually ignored an attribute because the task was too complex, while others may have ignored it because they do not derive utility from that attribute. In the first case (too complex), a person who ignored an attribute may have the same preferences as someone who attended to it and thus may also have the same WTP. In this case, the WTP based on the considered attributes would be applicable for everyone. In the second case, where respondents ignore an attribute because they do not derive any utility from it, the WTP for the ignored group is zero. Because we do not know the actual reasons for visually ignoring the attributes, the WTP estimates are calculated only using the

coefficient estimates for the considered attributes from Table 5, resulting in WTP estimates for sustainability labels of \$1.73¹² or lower. The Krinsky and Robb method (Krinsky and Robb, 1986; 1990) indicated no statistical differences in the mean estimates for the WTP distribution between the full attendance model and the models accounting for visual ANA. This result suggests that accounting for visual ANA, in our study, did not have significant implications in terms of mean WTP estimates.

While coffee prices largely depend on the quality of the beans, their origin, the blend, and the brand, organic coffee prices are often higher than prices for coffee certified by the Rainforest Alliance or Fair Trade USA (FAO, 2009). Retail prices for ground coffee ranged from \$3.00 to \$9.99 per 12-ounce package (store check, *City, state, 2013*) (*omitted for refereeing purposes only*). Because the prices for both conventional and certified coffee vary considerably by the quality and origin as well as by the nature of the outlet and brand, it is very difficult to determine a premium that is solely attributable to a sustainability certification rather than these other factors (Consumers International, 2005; FAO, 2009). Based on a store check (*City, state, 2013*) (*omitted for refereeing purposes only*), the price premium for coffee with a sustainability label ranges from \$1.50 to \$2.30 per 12 ounces when comparing coffees with and without a label of the same brand.

4 Conclusion

ANA is an important methodological issue for researchers engaged in CEs. Researchers cannot assume that respondents have attended to all the attributes in a CE and have processed all of the information in a fully rational manner. Assuming that a person considered an attribute when he or she actually ignored it could result in a violation of the continuity axiom, which may then lead to biased and misleading estimates. Given that CEs are commonly used to assess attribute

¹² This corresponds with $\widehat{WTP}_{organic} = \frac{\widehat{\beta}_{organic}}{-\widehat{\beta}_{price}} = 1.407/-0.812$ for model *defA-S*

valuation, there is an urgent need for and considerable research interest in methods to account for ANA.

As discussed previously, the two approaches that have emerged to address ANA are the stated and inferred approaches. The stated ANA approach relies on self-reported measures, while the inferred ANA approach relies on latent models or on derived respondent-specific coefficients. Thus, these approaches do not directly measure which attributes are actually being ignored in a CE. Additionally, evidence suggests that people do not consistently use the same attribute attendance rule throughout an entire sequence of choice tasks. As such, it is important to obtain ANA information at the choice task level. However, obtaining such information is difficult with the stated approach because asking about ANA in one choice task may trigger changes in attribute attendance in subsequent choice tasks. Additionally, making the CE longer could induce non-attendance. The inferred ANA approach also assumes that respondents ignore an attribute throughout an entire CE. Therefore, there is an urgent need to study other ways to address respondents' true processing strategies and attendance at the choice task level.

We contribute to this research area by using eye-tracking measures to evaluate visual attendance to the attributes in a CE. This method does not rely on self-reported ANA behaviour and does not attempt to infer ANA based on respondents' choice behaviour. Instead, we measured participants' visual attention to specify whether they visually attended to the information presented to them and extend the work of Balcombe et al. (2015) on visual ANA in CEs using eye-tracking technology in three ways. We use two fixation count cut-offs (i.e., one and two). Furthermore, we compare two definitions for detecting whether a specific attribute was ignored during a choice task based on the visual attention paid to the specific attribute in the choice set as a whole and to the specific attribute in each of the alternatives within the choice set. Additionally, we use two modelling methods to account for visual ANA: one at the choice set level (choice task visual ANA) and one at the respondent level (serial visual ANA).

Consequently, we use a total of six combinations or approaches to account for visual ANA based on the two modelling approaches, the two definitions, and two fixation count cut-offs, while Balcombe et al. (2015) used one combination, namely, the definition based on the visual attention paid to the attributes summated over the alternatives in the choice set, the serial ANA modelling approach, and a fixation count cut-off of two.

For each combination of fixation cut-off, visual ANA definition and modelling approach, we identified visually ignored attributes. Next, the coefficients of ignored attributes were restricted to zero which led to a decrease in model fit compared to the full attendance model. In addition, accounting for visual ANA did not significantly influence the WTP estimates. We then tested whether these attributes that were classified as ‘being ignored’ based on eye-tracking were truly ignored based on the respondents’ choices by estimating the coefficients of the ignored attributes. Interestingly, we found that not all visually ignored attributes were truly ignored based on the respondents’ choice behaviours (price and USDA Organic labels in most cases and Rainforest Alliance in some), i.e., the respondents paid some attention to these attributes in their choice consideration. Hence, our results suggest that Fair Trade, Carbon Footprint and in most cases Rainforest Alliance were truly ignored when they were identified as ignored using visual ANA. However, the same does not hold for price and USDA Organic. This result shows that the best way to account for visual ANA might depend on the attribute itself. For price and USDA Organic, it is likely that people do not need to fixate their eyes on these attributes to the same extent as the other attributes because it is possible that they are already acquainted or more familiar with these attributes. It is possible that for USDA Organic and price, they were aware of the information presented from other parts of their visual field (parafoveal and peripheral vision) not covered by eye fixations.

The price, USDA Organic, and (in some cases) Rainforest Alliance attributes that may have been classified as visually non-attended were in fact not truly ignored. Most coefficient

estimates for attributes that were classified as visually ignored were lower than the considered coefficients. This result might occur for two reasons. First, respondents who paid less attention to these attributes received less negative (price) or positive (USDA Organic, Rainforest Alliance) utility from these attributes. Second, the smaller coefficient of the ignored subset might be a combination of truly ignored attributes (zeros) and attended attributes and therefore result in a lower average. The true explanation is likely to be a combination of both reasons.

Next, the allocation of attenders and non-attenders based on visual ANA was compared with the allocation based on the coefficient of variation, an inferred method. The method of the coefficient of variation indicated that a small proportion of those classified as having ignored an attribute using the eye-tracking's visual ANA were also identified as having ignored it based on the inferred method. Given the insignificance of many of the coefficients of the ignored attributes, the mismatch between the allocation based on the inferred method and the serial visual ANA and the decrease in the model fit when using visual ANA, we can conclude, at least based on our study, that using eye-tracking does not always provide an adequate measure that can guarantee that an attribute should be considered as having been ignored or not.

Nevertheless, eye-tracking measures can provide useful information regarding respondents' decision-making behaviour. It can help us to better understand their decision-making process and its relation with choice, preferences, and attention (Greibitus et al., 2015). Previous studies illustrated that eye-tracking is useful to quantify visual attention, a measure for the degree of attention, during food choices (Van Loo et al., 2015; Lewis et al., 2016); however Balcombe et al. (2017) report that the relationship between visual attention and preference is weak. While eye-tracking has been used to quantify visual attention during food choice, applying it to identify visual ANA is more challenging, as suggested by our findings. We discuss these challenges and limitations in the following sections. Specifically, we first discuss

the limitations of eye-tracking research in general, followed by limitations specifically related to ANA research.

5 Limitations and challenges

5.1 The use of eye-tracking

Eye-tracking provides measures of where participants fixate but not why; thus, the motivations and cognitions underlying these eye movements remain unknown (Graham et al., 2012). Familiarity might influence how extensively a piece of information is examined, as possibly observed with the price and USDA Organic attributes (Pieters et al., 1996, 1999; Graham et al., 2012). Moreover, participants' fixations do not necessarily imply understanding and do not reveal anything about the higher-level processes of attention and comprehension. As suggested by Graham et al. (2012), conducting an interview after an eye-tracking task may provide more insight into what respondents were thinking during the task. While eye-tracking studies might be less prone to social desirability compared to studies that ask respondents directly about the information to which they attend, knowing that their eye movements will be monitored could potentially influence their behaviour (Graham et al., 2012). Finally, eye-tracking is a relatively expensive and time-consuming method. Hence, considerations with respect to value-for-money as well as budget and time constraints are also important.

5.2 The use of eye-tracking to study ANA

While eye-tracking monitors consumers' visual attention, this technology has some limitations specifically when studying visual ANA; i.e., when using it to determine whether a piece of information or an attribute was actually ignored or not, including those related to 1) the use of fixations, 2) the size of the AOIs, 3) the definitions for visual ANA, and 4) the influence of familiarity.

First, eye-tracking uses fixations as an indication of visual attention, but it only reveals what happens in an individual's foveal vision. While fixations take place in our foveal vision,

which is where our primary attention is focused (Bergstrom and Shall, 2014), eye-tracking does not allow us to measure the attention paid in one's parafoveal and peripheral vision (Bergstrom and Shall, 2014; Orquin et al., 2016). If the participants did not fixate on information presented in an AOI, it does not necessarily mean that they were not aware that it was there (Bergstrom and Shall, 2014, p 6). They may still have grasped information present in other parts of their visual field without focusing on it. While eye fixations within a certain AOI are believed to provide reliable measures for visual attention for that AOI (Bialkova and van Trijp, 2011), eye fixations do not necessarily represent everything that participants might have observed, making it challenging to use it as a measure for visual ANA, i.e., identifying if it was visually ignored or not. Based on our study, we can conclude that fixations, which indicate where the attention is focused, do not necessarily allow us to identify whether the information was ignored.

Second, eye-tracking uses AOIs for which specific metrics (such as fixation counts) are calculated, and thus the results depend on how these areas are defined. Unfortunately, there is no consensus in the decision-making research literature on the definition of AOIs. As mentioned by Orquin et al. (2016, p 103), 'this lack of standardization in AOI definition and reporting presents direct problems to the advancement of behavioural decision-making research'. Thus, there is a need to have a more standardized way to define AOIs, as the size of the AOI can influence its fixation count as well as other eye-tracking metrics in use.

Third, there are different ways to define visual ANA with respect to the number of fixation counts and the number of choice tasks. Balcombe et al. (2015) used the threshold of at least two fixations to consider an attribute as attended. In addition, they identify a respondent as a non-attender when such respondent ignores an attribute in more than half of the choice tasks. These are two rather arbitrary values and will influence the discrete measure of visual non-attendance (ignored versus not ignored).

Fourth, how familiar respondents are with the attribute or logo – as in our study – might have an influence. If they are very familiar with it, they may not need to fixate on it to recognize that the logo is present. If they are less familiar, they may need to look longer at and fixate more on the information for it to be processed. Thus, the decisions about the fixation cut-off as well as the size of the AOI for identifying information as being ignored might depend on the respondent's familiarity.

The four limitations discussed above represent major challenges, but they can hopefully assist researchers and future studies to further optimize the use of eye-tracking in the context of ANA in CEs. We discuss several specific suggestions for future research in the following subsection.

5.3 Suggestions for future research on the use of eye-tracking in CEs

While we attempted to account for the limitations described above as much as possible in the present study, there are several suggestions based on our study that future studies could take into account. In our study, the logos were rather large and in close proximity to each other. In addition, the coffee packages did not contain much other information. Since packages never had the same labels, it is possible that respondents could infer the quality attributes on one package based on the labels on the other package in a choice set. We recommend that future studies include other information cues on the food packages (e.g., nutritional labels, ingredients lists, graphics) that compete for the respondent's attention and make it more difficult for the participant to be aware of the presence of certain logos without fixating on them.

In this study, we used a head-free, below-screen eye-tracker that provided us with a relatively high tracking ratio (89.5%) across our respondents. Previous research has shown no significant difference between head-mounted and head-free types of eye trackers with respect to the pupil size-based index of cognitive activity (Bartels and Marshall, 2012). Furthermore, the

head-free, below-screen tracker allows participants to be less constrained and may promote more natural human behaviour. Nevertheless, future studies should also consider using a chin-rest, since this could avoid movement of the head, possibly resulting in better tracking quality (in terms of reliability and accuracy) compared with the head-free, below-screen tracker that we used in this study.

While we compared different visual ANA approaches, future studies could expand on comparing various ANA approaches (visual, stated, and inferred) and evaluating which of these techniques or combinations of techniques is the most appropriate to account for ANA. Specifically, for visual ANA, future research should improve and standardize how to apply visual ANA in choice modelling based on eye-tracking and remove any arbitrary steps. Studies could, for example, test other fixation count cut-offs and other visual ANA definitions. For example, for serial ANA, we used the approach used by Balcombe et al. (2015), whereby a respondent is identified as a non-attender if he/she ignored an attribute in more than half of the choice tasks. However, ‘more than half’ is an arbitrary approach, and other approaches could be tested to further fine-tune the serial visual ANA method. In addition, other AOI sizes and margins could be tested. Findings on these aspects could assist in optimizing the use of eye-tracking to address ANA and could provide additional guidelines on how to incorporate visual ANA into choice modelling research.

In addition to incorporating ANA or other decision-making strategies into choice models, future studies could use eye-tracking as a tool to study how to optimize experimental designs that would discourage respondents from ignoring information, thereby avoiding violation of the assumption of rational utility maximization (i.e., that the complete information presented is considered by respondents). We share the opinion of Balcombe et al. (2015), who suggested the use of eye-tracking to assist in the visual design and appearance of the CE instrument. Several eye-tracking studies have demonstrated that bottom-up factors can trigger

attention (Bialkova and van Trijp, 2011; Visschers et al., 2010). Bottom-up or stimulus-driven forms of attention are caused by characteristics of the visual stimulus itself (colour, size, location, saliency) and occur without specifically searching for them (Wolfe, 1998). Further insights into these bottom-up factors can help in the visual design or layout of choice sets that can reduce visual ANA in CEs. Spinks and Mortimer (2016) for example, investigated the impact of CE complexity (in terms of number of attributes ranging from 3 to 8) on the prevalence of ANA. They reported higher levels of ANA when more attributes are present and provided information on how to improve the CE design measuring health care preferences. They suggest the use of eye-tracking during the CE design and piloting phase, which is also suggested by Balcombe et al. (2017). Future studies could test other issues of the visual CE design and its impact on ANA such as the pictorial or verbal representation of attributes or alternatives, and other variations in representation of choice sets that could reduce ANA in CE studies.

Finally, in addition to ANA research, eye-tracking can be used to study other decision-making strategies and decision heuristics and can assist to explain the decision-making process in relation to food choices (Uggeldahl et al., 2016) and preferences. As mentioned by Orquin and Loose (2013), more research on how to effectively combine attention research from eye-tracking with decision-making research is warranted.

Acknowledgements

The authors thank the editor (Iain Fraser) and three anonymous journal reviewers for their valuable comments and suggestions that helped us to significantly improve the quality of the paper.

References

- Alemu, M.H., Mørkbak, M.R., Olsen, S.B. and Jensen, C.L. (2013). Attending to the reasons for attribute non-attendance in choice experiments. *Environmental & Resource Economics* 54:333-359.
- Balcombe, K., Burton, M. and Rigby, D. (2011). Skew and attribute non-attendance within the Bayesian mixed logit model. *Journal of Environmental Economics and Management* 62:446-461.
- Balcombe, K. G., Fraser, I. M. and McSorley, E. (2015). Visual attention and attribute attendance in multi-attribute choice experiments. *Journal of Applied Econometrics* 30:447–467.
- Balcombe, K., Fraser, I., Williams, L., and McSorley, E. (2017). Examining the relationship between visual attention and stated preferences: A discrete choice experiment using eye-tracking. *Journal of Economic Behavior & Organization* 144: 238-257.
- Barreiro-Hurle, J., Gracia, A. and de-Magistris, T. (2010). The effects of multiple health and nutrition labels on consumer food choices. *Journal of Agricultural Economics* 61:426-443.
- Bartels, M. and Marshall, S. P. (2012). Measuring cognitive workload across different eye tracking hardware platforms. 2012 ACM Symposium on Eye Tracking and Applications, Santa Barbara, California, 28-30 March, 2012.
- Bergstrom, J. R. and Schall, A. (2014). *Eye tracking in user experience design*. Waltham, MA: Elsevier.
- Bialkova, S. and van Trijp, H.C.M. (2011). An efficient methodology for assessing attention to and effect of nutrition information displayed front-of-pack. *Food Quality and Preference* 22:592-601.
- Bonnet, C. and Simioni, M. (2001). Assessing consumer response to protected designation of origin labelling: A mixed multinomial logit approach. *European Review of Agricultural Economics* 28:433-449.
- Burgess, L. (2007). Discrete Choice Experiments [computer software], Department of Mathematical Sciences, University of Technology, Sydney, available from <http://maths.science.uts.edu.au/maths/wiki/SPExpts>.
- Campbell, D., Boeri, M., Doherty, E. and Hutchinson, W.G. (2015). Learning, fatigue and preference formation in discrete choice experiments. *Journal of Economic Behavior and Organization* 119:345-363.
- Campbell, D., Hutchinson, W.G. and Scarpa, R. (2008). Incorporating discontinuous preferences into the analysis of discrete choice experiments. *Environmental & Resource Economics* 41:401-417.

- Campbell, D. and Lorimer, V.S. (2009). Accommodating attribute processing strategies in stated choice analysis: Do respondents do what they say they do. Amsterdam: 17th annual conference of the European Association of Environmental and Resource Economics.
- Campbell, D., Hensher, D.A. and Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: A latent class specification. *Journal of Environmental Planning and Management* 54:1061-1076.
- Caputo, V., Van Loo, E. J., Scarpa, R., Nayga, R. M., and Verbeke, W. (2017). Comparing serial, and choice task stated and inferred attribute non- attendance methods in food choice experiments. *Journal of Agricultural Economics* DOI: 10.1111/1477-9552.12246.
- Caputo, V., Nayga, R.M. and Scarpa, R. (2013). Food miles or carbon emissions? Exploring labelling preference for food transport footprint with a stated choice study. *Australian Journal of Agricultural and Resource Economics* 57:465-482.
- Carlsson, F., Kataria, M. and Lampi, E. (2010). Dealing with ignored attributes in choice experiments on valuation of Sweden's environmental quality objectives. *Environmental & Resource Economics* 47:65-89.
- Consumers International (2005). From bean to cup: How consumer choice impacts on coffee producers and the environment. London: Consumers International.
- Fairtrade International (2015). Monitoring the Scope and Benefits of Fair Trade, 7th ed. Bonn: Fairtrade International.
- Food and Agriculture Organization of the United Nations (FAO) (2009). *The Market for Organic and Fair-Trade Coffee*. Study prepared in the framework of FAO project GCP/RAF/404/GER: “Increasing Incomes and Food Security of Small Farmers in West and Central Africa Through Exports of Organic and Fair-Trade Tropical Products”. Trade and Markets Division. Rome: FAO.
- Graham, D. J., Orquin, J. L. and Visschers, V. H. (2012). Eye tracking and nutrition label use: A review of the literature and recommendations for label enhancement. *Food Policy* 37:378-382.
- Gracia, A., Barreiro-Hurle, J. and Perez, L.P.Y. (2012). Can renewable energy be financed with higher electricity prices? Evidence from a Spanish region. *Energy Policy* 50:784-794.
- Gracia, A. (2014). Consumers' preferences for a local food product: A real choice experiment. *Empirical Economics* 47:111-128.
- Grebitus, C., Roosen, J. and Seitz, C. C. (2015). Visual attention and choice: A behavioral economics perspective on food decisions. *Journal of Agricultural & Food Industrial Organization* 13:73-81.

- Greene, W. H. (1990). LIMDEP version 5.0. New York: Econometric Software. Inc.
- Greene, W. (2012). *Nlogit Version 5.0: Reference Guide*. Plainview, NY: Econometric Software.
- Grunert, K.G. (2011). Sustainability in the food sector: A consumer behaviour perspective. *International Journal on Food System Dynamics* 2:207-218.
- Henderson, J. M. and Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science* 10:438-443.
- Henderson, J. M., Williams, C. C., Castelano, M. S. and Falk, R. J. (2003). Eye movements and picture processing during recognition. *Perception & Psychophysics* 65:725-734.
- 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. (2014). Process heuristics in choice analysis: An editorial. *Journal of Choice Modelling* 11:1-3.
- Hensher, D.A., Rose, J.M. and Greene, W.H. (2005). The implications on willingness to pay of respondents ignoring specific attributes. *Transportation* 32:203-222.
- Hensher, D.A. and Rose, J.M. (2009). Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. *Transportation Research Part E* 45:583-590.
- Hensher, D.A. and Greene, W.H. (2010). Non-attendance and dual processing of common-metric attributes in choice analysis: A latent class specification. *Empirical Economics* 39:413-426.
- Hensher, D.A., Rose, J.M. and Greene, W.H. (2012). Inferring attribute non-attendance from stated choice data: Implications for willingness to pay estimates and a warning for stated choice experiment design. *Transportation* 39:235-245.
- Hensher, D.A., Rose, J.M. and Greene, W.H. (2015). *Applied choice analysis*. Second edition. Cambridge, UK: Cambridge University Press.
- Hess, S. and Hensher, D.A. (2010). Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. *Transportation Research Part B* 44:781-790.
- Hess, S., Stathopoulos, A., Campbell, D., O'Neill, V. and Caussade, S. (2013). It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity. *Transportation* 40:583-607.
- Hole, A.R. (2011). A discrete choice model with endogenous attribute attendance. *Economics Letters* 110:203-205.

- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. and Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. *Ecological Economics* 69:1595-1603.
- Kehlbacher, A., Balcombe, K. and Bennett, R. (2013). Stated attribute non-attendance in successive choice experiments. *Journal of Agricultural Economics* 64:693–706.
- Kragt, M.E. (2013). Stated and inferred attribute attendance models: A comparison with environmental choice experiments. *Journal of Agricultural Economics* 64:719-736.
- Krinsky, I. and Robb, A. L. (1986). On approximating the statistical properties of elasticities. *The Review of Economics and Statistics* 68:715-719.
- Krinsky, I. and Robb A. L. (1990). On approximating the statistical properties of elasticities: A correction. *The Review of Economics and Statistics* 72:189-190.
- Krucien, N., Ryan, M., and Hermens, F. (2017). Visual attention in multi-attributes choices: What can eye-tracking tell us?. *Journal of Economic Behavior & Organization* 135:251-267.
- Lagarde, M. (2013). Investigating attribute non-attendance and its consequences in choice experiments with latent class models. *Health Economics* 22:554-567.
- Lancaster, K.J. (1966). New approach to consumer theory. *Journal of Political Economy* 74:132-157.
- Lancsar, E. and Louviere, J. (2006). Deleting ‘irrational’ responses from discrete choice experiments: A case of investigating or imposing preferences? *Health Economics* 15:797-811.
- Layton, D.F. and Brown, G. (2000). Heterogeneous preferences regarding global climate change. *The Review of Economics and Statistics* 82:616–24.
- Lewis, K. E., Grebitus, C. and Nayga, R. M. (2016). The impact of brand and attention on consumers’ willingness to pay: Evidence from an eye tracking experiment. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 64:753-777.
- Lusk, J. L. and Schroeder, T. C. (2004). Are choice experiments incentive compatible? A test with quality differentiated beef steaks. *American Journal of Agricultural Economics* 86:467-482.
- Mariel, P., Hoyos, D. and Meyerhoff, J. (2013). Stated or inferred attribute non-attendance? A simulation approach. *Economía Agraria y Recursos Naturales (Agricultural and Resource Economics)* 13:51-67.

- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P. (ed), *Frontiers in Econometrics*. New York: Academic Press, 105-142.
- Meyerhoff, J. and Liebe, U. (2009). Discontinuous preferences in choice experiments: Evidence at the choice task level. Amsterdam: 17th annual conference of the European Association of Environmental and Resource Economics.
- Mintel. (2012). *Coffee – USA - October 2012*. London: Mintel Group Ltd.
- Mørkbak, M.R., Olsen, S.B. and Campbell, D. (2014). Behavioral implications of providing real incentives in stated choice experiments. *Journal of Economic Psychology* 45:102-116.
- Morey, E. and Rossmann, K. (2003). Using stated-preference questions to investigate variations in willingness to pay for preserving marble monuments: Classic heterogeneity, random parameters, and mixture models. *Journal of Cultural Economics* 27:215-29.
- Orquin, J. L., Ashby, N. J. and Clarke, A. D. (2016). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making* 29:103-115.
- Orquin, J. L. and Loose, S. M. (2013). Attention and choice: A review on eye movements in decision making. *Acta psychological* 144: 190-206.
- Pierrot, J., Giovannucci, D., and Kasterine, A. (2011). Trends in the trade of certified coffees. Technical Paper. Geneva: International Trade Centre (ITC).
- Pieters, R.G.M., Rosbergen, E. and Hartog, M. (1996). Visual attention to advertising: The impact of motivation and repetition. *Advances in Consumer Research* 23:242–248.
- Pieters, R.G.M., Rosbergen, E. and Wedel, M. (1999). Visual attention to repeated print advertising: A test of scanpath theory. *Journal of Marketing Research* 36:424–438.
- Poole, A., Ball, L. and Phillips, P. (2005). In search of salience: A response-time and eye-movement analysis of bookmark recognition. In Fincher, S., Markopoulos, P., Moore, D. and Ruddle, R. (eds), *People and Computers Xviii — Design for Life*. London: Springer, 363-378.
- Puckett, S.M. and Hensher, D.A. (2008). The role of attribute processing strategies in estimating the preferences of road freight stakeholders. *Transportation Research Part E* 44:379-395.
- Puckett, S.M. and Hensher, D.A. (2009). Revealing the extent of process heterogeneity in choice analysis: An empirical assessment. *Transportation Research Part A* 43:117-126.
- Revelt, D. and Train, K. (1998). Mixed logit with repeated choices: Households' choices of appliance efficiency level. *The Review of Economics and Statistics* 80:647–57.
- Ryan, M. and Bate, A. (2001). Testing the assumptions of rationality, continuity and symmetry when applying discrete choice experiments in health care. *Applied Economics Letters* 8:59-63.

- Scarpa, R., Ferrini, S. and Willis, K. (2005). Performance of error component models for status-quo effects in choice experiments. In Scarpa, R. and Alberini, A. (eds.), *Applications of Simulation Methods in Environmental and Resource Economics*. Dordrecht: Springer, 247-273.
- Scarpa, R., Campbell, D. and Hutchinson, W.G. (2007). Benefit estimates for landscape improvements: Sequential Bayesian design and respondents' rationality in a choice experiment. *Land Economics* 83:617-634.
- Scarpa, R., Gilbride, T.J., Campbell, D. and Hensher, D.A. (2009). Modelling attribute non-attendance in choice experiments for rural landscape valuation. *European Review of Agricultural Economics* 36:151-174.
- 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* 86:817-839.
- Scarpa, R., Zanolli, R., Bruschi, V. and Naspetti, S. (2013). Inferred and stated attribute non-attendance in food choice experiments. *American Journal of Agricultural Economics* 95:165-180.
- SMI (SensoMotoric Instruments). (2016). BeGaze Manual. Version 3.6. February 2016. 416 pages.
- Spinks, J., and Mortimer, D. (2016). Lost in the crowd? Using eye-tracking to investigate the effect of complexity on attribute non-attendance in discrete choice experiments. *BMC Medical Informatics and Decision Making* 16(1):14.
- Street, D.J. and Burgess, L. (2007). *The Construction of Optimal Stated Choice Experiments*. Hoboken: John Wiley & Sons, Inc.
- Uggeldahl, K., Jacobsen, C., Lundhede, T. H., and Olsen, S. B. (2016). Choice certainty in discrete choice experiments: Will eye tracking provide useful measures?. *Journal of choice modelling* 20:35-48.
- Van Loo, E. J., Caputo, V., Nayga Jr, R.M. and Verbeke, W. (2014). Consumers' valuation of sustainability labels on meat. *Food Policy* 49:137-150.
- Van Loo, E.J., Caputo, V., Nayga Jr., R. M., Seo, H.-S., Zhang, B. and Verbeke, W. (2015). Sustainability labels on coffee: Consumer preferences, willingness-to-pay and visual attention to attributes. *Ecological Economics* 118:215-225.
- Vischers, V.H.M., Hess, R. and Siegrist, M. (2010). Health motivation and product design determine consumers' visual attention to nutrition information on food products. *Public Health Nutrition* 13:1099-1106.

Wolfe, J. M. (1998). Visual search. In Pashler, H. (Ed.), *Attention*. East Sussex, UK: Psychology Press, 13-22.

Figure 1 shows a choice set question with three coffee options. Each option is represented by a coffee bag image and a price tag. The first option is a 12 oz (340g) bag of Ground Medium Roast coffee, priced at \$6.30. It features a USDA Organic label, a Carbon Footprint label, and a Rainforest Alliance label. The second option is a 12 oz (340g) bag of Ground Medium Roast coffee, priced at \$8.30. It features a Fair Trade label, a Rainforest Alliance label, and a Rainforest Alliance label. The third option is 'None of these'. Below the options are three empty boxes for selection.

Please select the option you would choose in the store. (click one time in the box of your choice)
Please select <SPACE> when you are done.

Figure 1. Example of choice set question with AOI. Frames indicate the AOI for the Carbon Footprint label, Rainforest Alliance label, Fair Trade label, USDA Organic label and price.

Table 1. Socio-demographic characteristics of the sample (% , n = 81).

Gender	
Male	46.9
Female	53.1
Age group	
18-24 years	17.3
25-34 years	37.0
35-44 years	21.0
45-54 years	14.8
55 and older	9.9
Children	
Yes	54.3
No	45.7
Educational level completed	
High school/GED or lower	7.4
Some college or 2-year college degree	27.2
4-year college degree (BA, BS)	39.5
Master's or PhD degree	25.9
Annual household income	
Less than \$20,000	25.9
\$20,000-\$49,999	35.8
\$50,000-\$99,999	25.9
More than \$100,000	12.4

Table 2. Attributes and levels used in the choice experiment

Attributes	Level
Fair Trade label	
USDA Organic label	0 = Not present
Rainforest Alliance label	1 = Present
Carbon Footprint label	
	0 = \$4.30
	1 = \$6.30
Price (per 12 ounces)	2 = \$8.30
	3 = \$10.30

Table 3. Overview of the six combinations based on the two modelling approaches to account for visual ANA, the fixation count used as the cut-off, and the two definitions for detecting whether a specific attribute was ignored during a choice task.

Fixation count		FC 2				FC 1	
ANA modelling approach	Full attendance	Serial ANA		Choice task ANA		Serial ANA	Choice task ANA
Definition ignored attribute		Def. A Based on choice set	Def. B Based on alternatives	Def. A Based on choice set	Def. B Based on alternatives		
Abbreviation	<i>Full-AA</i>	<i>defA-S</i>	<i>defB-S</i>	<i>defA-CT</i>	<i>defB-CT</i>	<i>FC1-S</i>	<i>FC1-CT</i>

Table 4. Proportions (%) of choice task and serial visual ANA depending on the definition applied and fixation count (n=645)

	FC = 2				FC = 1	
	Serial ANA		Choice task ANA		Serial ANA	Choice task ANA
	Def A	Def B	Def A	Def B		
Fair Trade	40.7	40.7	52.2	52.2	12.3	28.1
USDA Organic	55.6	55.6	59.6	59.6	23.5	30.4
Rainforest Alliance	43.2	43.2	54.9	54.9	17.3	30.1
Carbon Footprint	40.7	40.7	52.9	52.9	16.0	28.1
Price	12.3	23.4	24.5	40.7	4.9	9.6

Table 5. RPL model with error component (RPL-EC) parameter estimates (n=645)

Fixation count		FC 2										FC 1			
ANA modelling		Full AA		Serial ANA		Choice task ANA						Serial ANA		Choice task ANA	
Def. ignored attribute				Def. A		Def. B		Def. A		Def. B					
				<i>defA-S</i>		<i>defB-S</i>		<i>defA-CT</i>		<i>defB-CT</i>		<i>FC1-S</i>		<i>FC1-CT</i>	
		Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.
Fair Trade	Mean	0.63**	(0.29)	0.73**	(0.36)	0.75**	(0.33)	0.76***	(0.27)	0.81***	(0.22)	0.65**	(0.31)	0.65**	(0.29)
	St. Dev.	0.83***	(0.28)	0.79**	(0.32)	0.73**	(0.37)	0.65	(0.51)	0.54	(0.47)	0.58*	(0.29)	0.40	(0.30)
USDA Organic	Mean	1.02***	(0.35)	1.41***	(0.50)	1.35***	(0.47)	0.87***	(0.25)	0.84***	(0.24)	1.09***	(0.40)	0.77***	(0.26)
	St. Dev.	1.23***	(0.33)	0.85*	(0.52)	0.89	(1.05)	0.39	(0.49)	0.18	(1.14)	1.13**	(0.50)	0.84*	(0.47)
Rainforest Alliance	Mean	0.74***	(0.26)	0.73***	(0.26)	0.68***	(0.23)	0.60***	(0.23)	0.60***	(0.21)	0.75***	(0.27)	0.59***	(0.20)
	St. Dev.	0.58	(0.37)	0.45	(0.41)	0.33	(0.44)	0.42	(0.73)	0.15	(1.60)	0.55	(0.52)	0.36	(0.52)
Carbon Footprint	Mean	0.30	(0.26)	0.38	(0.36)	0.40	(0.32)	0.50**	(0.23)	0.58***	(0.22)	0.27	(0.31)	0.34	(0.22)
	St. Dev.	0.90*	(0.48)	0.72	(0.74)	0.71	(0.98)	0.39	(0.77)	0.39	(1.27)	0.62	(0.50)	0.39	(1.03)
Price		-0.85***	(0.05)	-0.81***	(0.05)	-0.84***	(0.05)	-0.61***	(0.05)	-0.54***	(0.06)	-0.77***	(0.05)	-0.68***	(0.04)
No_Buy		-8.76***	(0.86)	-8.91***	(0.85)	-9.13***	(0.98)	-7.35***	(0.90)	-6.75***	(1.01)	-8.01***	(0.81)	-7.61***	(0.81)
Err. Comp.	St. Dev.	2.71***	(0.91)	3.15***	(1.15)	3.63***	(0.85)	3.03***	(0.96)	3.51***	(0.93)	2.55***	(0.96)	2.63***	(0.90)
Log likelihood		-349		-354		-370		-404		-433		-357		-375	
AIC		731.2		741.2		774.1		842.4		900.6		747.9		783.6	
BIC		807.2		817.2		850.0		918.4		976.6		823.9		859.6	

Note: *, ** and *** indicate statistical significance at 10%, 5% and 1% levels, respectively.

Table 6. MNL parameter estimations with two coefficients (considered and ignored) (n=645)

Fixation count	FC 2								FC 1			
ANA modelling	Serial ANA				Choice task ANA				Serial ANA		Choice task ANA	
Def. ignored attribute	Def A		Def B		Def A		Def B		FC1-S2		FC1-CT2	
	<i>defA-S2</i>		<i>defB-S2</i>		<i>defA-CT2</i>		<i>defB-CT2</i>		<i>FC1-S2</i>		<i>FC1-CT2</i>	
	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.
<i>Considered</i>												
Fair Trade	0.54***	(0.16)	0.55***	(0.17)	0.72***	(0.18)	0.72***	(0.18)	0.43***	(0.14)	0.46***	(0.15)
USDA Organic	1.46***	(0.20)	1.45***	(0.20)	1.11***	(0.20)	1.12***	(0.20)	0.99***	(0.16)	0.94***	(0.16)
Rainforest Alliance	0.86***	(0.18)	0.87***	(0.18)	0.93***	(0.19)	0.93***	(0.19)	0.72***	(0.15)	0.73***	(0.16)
Carbon Footprint	0.16	(0.17)	0.18	(0.17)	0.37**	(0.18)	0.38**	(0.18)	0.06	(0.14)	0.17	(0.15)
Price	-0.65***	(0.05)	-0.65***	(0.05)	-0.62***	(0.05)	-0.63***	(0.05)	-0.61***	(0.05)	-0.62***	(0.05)
<i>Ignored</i>												
Fair Trade	0.10	(0.18)	0.08	(0.18)	0.03	(0.17)	0.03	(0.17)	-0.41	(0.33)	-0.03	(0.22)
USDA Organic	0.39**	(0.17)	0.40**	(0.17)	0.81***	(0.17)	0.80***	(0.17)	0.35	(0.24)	0.69***	(0.22)
Rainforest Alliance	0.30	(0.19)	0.30	(0.19)	0.38**	(0.17)	0.38**	(0.17)	0.08	(0.28)	0.34	(0.22)
Carbon Footprint	-0.01	(0.19)	-0.04	(0.19)	-0.13	(0.17)	-0.14	(0.17)	0.11	(0.28)	-0.17	(0.22)
Price	-0.40***	(0.08)	-0.52***	(0.06)	-0.51***	(0.06)	-0.54***	(0.05)	-0.46***	(0.12)	-0.43***	(0.08)
No_Buy	-5.47***	(0.40)	-5.38***	(0.40)	-5.09***	(0.39)	-5.06***	(0.39)	-5.28***	(0.39)	-5.18***	(0.39)
Log likelihood ¹	-376		-379		-386		-386		-388		-393	
AIC	774.1		779.8		793.1		794.5		797.1		807.0	
BIC	823.2		829.0		842.3		843.7		846.2		856.2	
<i>p-values for testing statistical differences between the coefficients considered versus ignored</i>												
Fair Trade	0.024**		0.016**		0.001***		0.001***		0.006***		0.018**	
USDA Organic	<0.001***		<0.001***		0.081*		0.074*		0.006***		0.144	
Rainforest Alliance	0.007***		0.005***		0.007***		0.007***		0.016**		0.051*	
Carbon Footprint	0.214		0.165		0.013**		0.010***		0.430		0.081*	
Price	0.001***		0.005***		0.008***		0.013**		0.088*		0.008***	

Note: *, ** and *** indicate statistical significance at 10%, 5% and 1% levels, respectively.

¹ For the Full AA MNL model: log likelihood = -401.4, AIC= 814.8, BIC=841.7 (see Appendix B).

Table 7. Comparison of allocation of respondents for serial visual ANA and serial inferred ANA (count) (n=81)

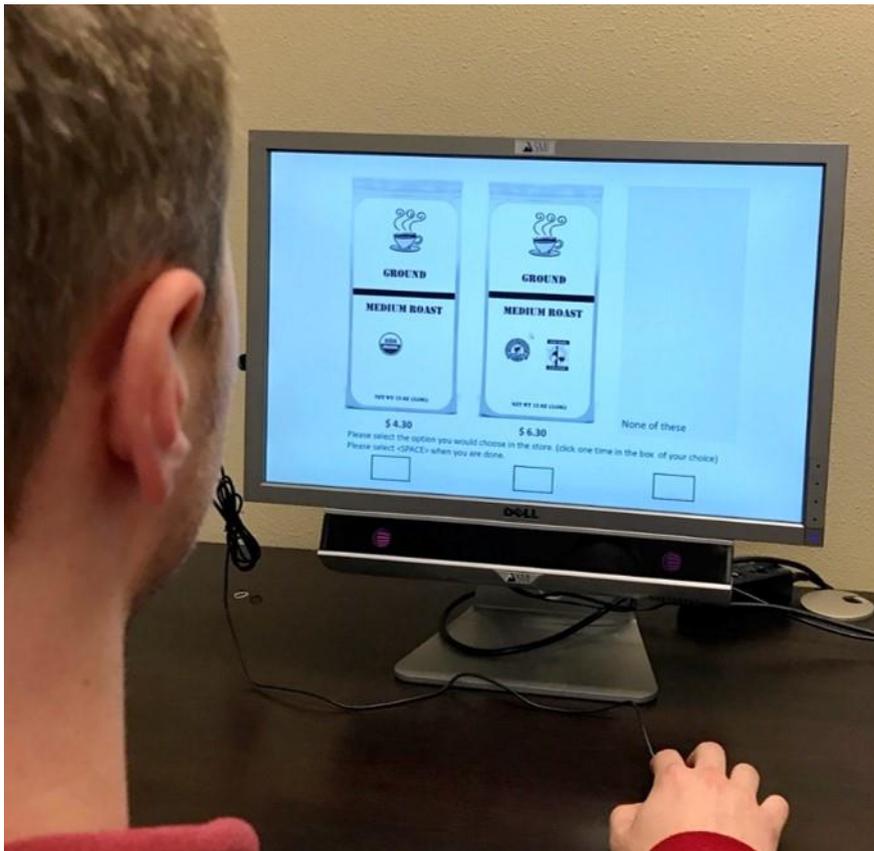
Serial inferred ANA		Serial visual ANA					
		FC 2			FC 1		
		NI	I	NI + I (%total)	NI	I	NI + I (%total)
Fair Trade	NI	41	25	66 (81.5%)	60	6	66 (81.5%)
	I	7	8	15 (18.5%) ¹	11	4	15 (18.5%)
	NI + I (%total)	48 (59.3%)	33 (40.7%) ¹		71 (87.7%)	10 (12.3%)	
Matching allocations (%)		60.5% ²			79.0%		
USDA Organic	NI	34	35	69 (85.2%)	56	13	69 (85.2%)
	I	2	10	12 (14.8%)	6	6	12 (14.8%)
	NI + I (%total)	36 (44.4%)	45 (55.6%)		62 (76.5%)	19 (23.5%)	
Matching allocations (%)		54.3%			76.5%		
Rainforest Alliance	NI	44	32	76 (93.8%)	63	13	76 (93.8%)
	I	2	3	5 (6.2%)	4	1	5 (6.2%)
	NI + I (%total)	46 (56.8%)	35 (43.2%)		67 (82.7%)	14 (17.3%)	
Matching allocations (%)		58.0%			79.0%		
Carbon Footprint	NI	37	15	52 (64.2%)	47	5	52 (64.2%)
	I	11	18	29 (35.8%)	21	8	29 (35.8%)
	NI + I (%total)	48 (59.3%)	33 (40.7%)		68 (84.0%)	13 (16.0%)	
Matching allocations (%)		67.9%			67.9%		

NI: Not ignored; I=Ignored

¹ Out of 81 respondents, 15 were considered as having ignored the Fair Trade attribute based on the inferred method and 33 were considered as having ignored the attribute based on serial visual ANA with fixation count 2.

² For 8 and 41 respondents, the classifications into ignoring and not ignoring, respectively, are the same between the inferred and visual methods, resulting in a matching allocation of 60% for fair trade $((41+8)/81)$.

1 Appendix A. Set-up of the eye-tracking experiment in the lab



2

3

4 Appendix B. MNL parameter estimates (n=645)

Fixation count	FC = 2						FC = 1							
	Full AA		Serial ANA		Choice task ANA		Serial ANA		Choice task ANA					
ANA modelling														
Def. ignored attribute			Def. A	Def. B	Def. A	Def. B								
			<i>defA-S</i>	<i>defB-S</i>	<i>defA-CT</i>	<i>defB-CT</i>	<i>FCI-S</i>		<i>FCI-CT</i>					
	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.	Coeff.	St. err.				
Fair Trade	0.32**	(0.13)	0.61***	(0.15)	0.66***	(0.14)	0.76***	(0.15)	0.80***	(0.14)	0.49***	(0.13)	0.57***	(0.13)
USDA Organic	0.81***	(0.14)	1.26***	(0.18)	1.08***	(0.16)	0.80***	(0.16)	0.75***	(0.15)	0.89***	(0.14)	0.74***	(0.14)
Rainforest Alliance	0.57***	(0.14)	0.68***	(0.15)	0.58***	(0.14)	0.60***	(0.15)	0.57***	(0.14)	0.62***	(0.14)	0.53***	(0.13)
Carbon Footprint	0.05	(0.13)	0.26*	(0.15)	0.33**	(0.14)	0.52***	(0.15)	0.59***	(0.14)	0.14	(0.13)	0.28**	(0.13)
Price	-0.59***	(0.05)	-0.61***	(0.05)	-0.51***	(0.04)	-0.49***	(0.04)	-0.38***	(0.04)	-0.58***	(0.05)	-0.57***	(0.04)
No_Buy	-5.12***	(0.38)	-5.36***	(0.38)	-4.45***	(0.34)	-4.35***	(0.34)	-3.35***	(0.30)	-5.10***	(0.37)	-4.99***	(0.36)
Log likelihood	-401.4		-394		-429		-444		-479		-397		-411	
AIC	814.8		800.5		870.0		899.7		969.4		805.8		833.4	
BIC	841.7		827.3		896.8		926.5		996.2		832.6		860.2	

5 Note: *, ** and *** indicate statistical significance at 10%, 5% and 1% levels, respectively.

6 ¹ With a fixation count of one as cut-off, definitions A and B are the same.

7 Appendix C. The Cholesky matrices corresponding to RPL models with error component (RPL-EC)
 8 from Table 5

9 Full-AA model

	Fair Trade	USDA Organic	Rainforest Alliance	Carbon Footprint
Fair Trade	0.83***			
USDA Organic	-0.28	1.19***		
Rainforest Alliance	-0.00	0.68	0.08	
Carbon Footprint	0.73	-0.45	-0.08	0.09

10

11 defA-S model

12

	Fair Trade	USDA Organic	Rainforest Alliance	Carbon Footprint
Fair Trade	0.79**			
USDA Organic	-0.67	0.01		
Rainforest Alliance	-0.28	0.31	0.11	
Carbon Footprint	0.51	-0.55	-0.17	0.02

13

14

15 defB-S model

	Fair Trade	USDA Organic	Rainforest Alliance	Carbon Footprint
Fair Trade	0.73**			
USDA Organic	-0.64	0.13		
Rainforest Alliance	-0.22	0.26	0.13	
Carbon Footprint	0.48	-0.61	-0.20	0.01

16

17 defA-CT model

	Fair Trade	USDA Organic	Rainforest Alliance	Carbon Footprint
Fair Trade	0.65			
USDA Organic	-0.23	0.17		
Rainforest Alliance	-0.25	0.10	0.01	
Carbon Footprint	0.25	-0.14	-0.14	0.05

18

19

20 defB-CT model

	Fair Trade	USDA Organic	Rainforest Alliance	Carbon Footprint
Fair Trade	0.54			
USDA Organic	-0.07	0.12		
Rainforest Alliance	-0.07	0.03	0.01	
Carbon Footprint	0.21	-0.04	-0.04	0.04

21

22 FC1-CT model

	Fair Trade	USDA Organic	Rainforest Alliance	Carbon Footprint
Fair Trade	0.40			
USDA Organic	-0.09	0.81**		
Rainforest Alliance	0.02	0.27	0.04	
Carbon Footprint	0.15	-0.17	-0.02	0.04

23

24 FC1-S model

	Fair Trade	USDA Organic	Rainforest Alliance	Carbon Footprint
Fair Trade	0.57*			
USDA Organic	-0.21	1.07***		
Rainforest Alliance	0.03	0.56	0.06	
Carbon Footprint	0.36	-0.27	0.01	0.02

25

26 Note: *, ** and *** indicate statistical significance at 10%, 5% and 1% levels, respectively.