The Potential Role of Eye Tracking in Stated Preference Survey Design and Piloting

William Brazil
Department of Civil, Structural and Environmental Engineering
Trinity College Dublin, Ireland
Tel: + 353 1 896 2534; Fax: + 353 1 677 3072
Email: brian.caulfield@tcd.ie

Brian Caulfield (corresponding author)
Department of Civil, Structural and Environmental Engineering
Trinity College Dublin, Ireland
Tel: + 353 1 896 2534; Fax: + 353 1 677 3072
Email: brian.caulfield@tcd.ie

Chandra R. Bhat
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin TX 78712, USA
Tel: 1-512-471-4535; Email: bhat@mail.utexas.edu
and
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

August 2017
ABSTRACT
Transport surveys are designed to provide information that can give greater insight into individuals’ mobility habits. It is important in these surveys that efforts be made to ensure that the data collected accurately reflects those of the subject population under examination. A stated preference survey is a particular format of surveys that has been of particular interest to transport researchers over the last few decades. These surveys provide researchers with a means of creating choice contexts that attempt to closely mirror real life choice contexts, thereby providing insight into the factors that influence individuals’ choices.

This paper outlines a stated preference experiment based upon the use of an image-based elicitation mechanism, designed to highlight the role that emerging low cost eye-tracking solutions can play to aid survey design. Specific attention is placed on the impacts of scenario layout and the prominence of alternatives and attribute levels on stated responses. The results demonstrate the complexity involved in information assimilation in stated preference scenarios, even in the relatively simple example used for this study.
INTRODUCTION
Transport is a sector in which the outcomes of any infrastructure/behavioral intervention or investment are highly linked to human behaviour. Transport planners make decisions based on stochastic (without complete certainty) predictions about how the public will react to new policies and plans. Such decisions are not, however, made in a vacuum; rather they are informed by a wide variety of tried and tested methods. Such methods are based upon experience, previous examples, and empirical data. Within such a context, a high degree of emphasis has been placed upon the role of mathematical models, which attempt to predict the behaviour of individuals based upon good theory and well established determinant factors. To inform such models, it is very important to have accurate underlying data, and one method of collecting data regarding potential behaviours is through transport surveys.

Stated preference (SP) surveys present participants with a number of hypothetical scenarios, and ask them to make a choice based on the information provided. Typically, this tends to involve the study of a specific choice context (such as travel mode choice), with multiple scenarios (also sometimes referred to as games or experiments). Using the responses of individuals to these choice scenarios, it is possible to estimate models that can then be used to compare the importance and impact of various attributes on individuals’ decision making, impute willingness to pay measures such as the value of travel time, and generate share predictions.

When developing the SP scenarios, the researcher attempts to design them in such a manner as to make them easily comprehensible to respondents, while also containing enough relevant information to create a realistic and operational model. The researcher may assume that survey respondents are weighting all the information they are presented with and making their decisions accordingly; however previous research, in the area of attribute non-attendance appears to show that this is not always the case (1; 2). The implication is that models may not be accurately reflecting underlying user motivations and likely future behaviors.

This paper presents the results of research undertaken using eye-tracking technology and SP travel scenarios to gain insights into the potential uses of eye-tracking technology in the survey design process, and to demonstrate the benefits of using this technology.

BACKGROUND
Stated Preference Scenario Design
Typically, SP surveys include objective factors as determinants of choice, such as travel time and trip cost in a travel mode choice context; however, these surveys can also be extended to include less quantitative factors, such as the comfort levels or perceived safety associated with an alternative. The analyst can then alter the values of these attributes, presenting respondents with scenarios that engender trade-offs between the attributes, in the hope of gaining a better understanding of the individual’s decision making process.

During SP experiments, it is assumed that the participants consider and assess all the information presented to them regarding the alternatives when making their choices. The responses can then be modelled using a variety of appropriate modelling methods to gain an understanding of the role that the attributes and their levels play in determining the preference for one option over another. But, as we discuss next, the assumption of comprehensive attendance to all attributes characterizing a choice scenario may not always be valid.
Attribute Non-Attendance
Evaluating a SP scenario can be quite a complex task for respondents, especially for those not familiar with SP surveys, leading participants to adopt strategies derived from heuristics to simplify the process (3). One such strategy is that of attribute exclusion or non-attendance. In essence this is when survey respondents fail to consider or consciously choose to ignore some information contained in one or many of the attributes. This can be in the form of totally ignoring the information, or only partially attending to the information (4). Non-attendance can be an issue for modellers too, as research has shown that non-attenders can have different preferences relative to attenders (5). This non-attendance is not a random occurrence, and can be related to factors such as knowledge, or lack thereof, of an area (6) suggesting a possible internal segmentation in the responses. Traditionally this issue has been addressed via supplementary questions or model specification (7; 8). However, recent advances in technology would suggest that there is the potential for technologies such as eye-tracking and virtual reality to help address this problem (3).

Testing and Piloting
As with any survey or data collection technique, SP surveys require extensive testing and piloting to ensure that they are meeting the requirements of the project, and are not creating unintentional biases within the responses. While eye-tracking technology cannot be considered to be a new development from a technological perspective, the emergence of low cost portable technology devices is greatly increasing the accessibility of such methods, ensuring that they are no longer confined to a small number of institutions and research groups. One result of these developments is the potential to apply this technology to survey design, both ensuring that participants are paying attention to the intended stimuli, and also as exploratory tools to pick up unforeseen behaviours.

APPLICATIONS OF EYE-TRACKING TECHNOLOGY
Eye-tracking is an established research technique in a number of fields such as marketing (9), road safety (10), hazard detection (11) and energy consumption (12). A number of studies have also been undertaken applying eye-tracking to methods to both survey design and specifically to stated choice research.

Eye-tracking in Choice Situations
While eye-tracking has traditionally been a niche area of research, there are a number of examples of where it has been applied to choice situations to provide a better insight into the decision making process. Oviedo and Caparrós (13) utilised eye-tracking to examine forestry in Spain, while Rasch et al (14) combined eye-tracking technology with facial electromyography in an examination of choices relating to yogurt and shower gels. Similarly, Campbell et al (15) used eye-tracking, in conjunction with a conjoint analysis experiment, to examine factors that impact plant retail sales. When examining the relationship between attention, as inferred from eye-tracking results, and the respondents’ choices based on three empirical studies, Meißner et al (16) found that the alternative that was examined first, as well as the alternative that was positioned centrally, received higher amounts of attention than the others on display. However, the authors stated that this additional attention given to an alternative does not reliably influence the probability of the respondent selecting that option. Looking specifically at the role of eye-tracking with regard to attribute non-attendance, Balcombe et al (17) showed that systematic non-attendance to attributes only occurs in a minority of cases, and that most respondents look at most of the attributes. This work also found that visual and stated attendance were informative and non-overlapping, suggesting that the eye-tracking technology could be complementary to other methods.
Eye-tracking and Information Assimilation

Eye-tracking has been used as a tool in the marketing sector to gain an understanding of how products compete for consumers’ attention, as well as to understand the most effective means of transmitting information to the consumer. When examining attribute-by-product matrices for comparison websites Shi et al (18) used eye-tracking to look at consumers’ information acquisition strategies. A number of studies have examined the role of labels as information dissemination tools using eye-tracking technology, specifically Siegrist et al (19) who examined elements of nutritional information, and Waechter et al (20) who assessed the effectiveness of EU energy labels as means of raising awareness of the energy impacts of household appliances.

Eye-tracking Applications in Transport

Eye-tracking has also been utilized in a number of studies, but not in SP or survey design. Typically, this technology is used in the field of road safety to assess the ability of transport users to identify hazards they may encounter on the road. Jones et al (21), in a study looking at the role of emotion and perception of risks in driving scenarios, used eye-tracking technology in conjunction with galvanic skin response and heart rate recordings. A similar study by Brazil et al et al (22) used Google Maps and eye-tracking technology to assess the relationship between cyclists’ hazard detection ability and their cycling experience. A study by Frings et al (23) used an eye-tracker mounted to a cycling helmet to examine attention allocation during junction negotiation, while Vansteenkiste et al (24) used eye-tracking in a study of how to improve cycle lane quality.

METHODOLOGY

Stated Preference Application

The eye-tracking technology was applied to a number of scenarios relating to a hypothetical intercity trip roughly 200-300 kilometres in length, which would reflect an average inter-city journey in the Republic of Ireland. Three attributes and three alternatives were selected for the purposes of this study. The model alternatives for the trip were a private car, an intercity coach, and a journey by mainline rail. The alternative attributes selected for examination were trip time, cost, and travel time variability. It was decided to only include quantitative attributes for ease of analysis in this initial exploration of eye-tracking technology for SP surveys.

Nine scenarios were created for this experiment, and the alternative label was represented both with the name of the alternative and with an image of the alternative. The scenarios were designed using an orthogonal design approach that created 27 scenarios in total, with the first eight that were generated being selected for the purposes of this analysis. While this study was not concerned with estimating discrete choice models, the orthogonal design approach was used to ensure that the scenarios were in line with those used in similar SP studies. An additional scenario, where the car option was designed to have worse-off values on each attribute, was included for the purposes of internal validation.

Two presentation methods were selected to allow for a comparison that would effectively demonstrate the abilities of the eye-tracking approach to compare scenario designs. The sample was therefore split, with half of respondents viewing an image of the mode of transport, and the other half only having its name outlined in text (see Figure 1). Figure 1 indicates, the attributes and attribute levels remain consistent between the presentation approaches, with the difference being only in the alternative representation.
**Experimental Setup**

The Eyetribe eye-tracker was used in conjunction with the accompanying Eye-Proof software (25) for our analysis. Participants were seated facing a laptop with the eye-tracking bar placed just below the screen (See Figure 2). The bar was angled in such a manner as to capture the movement of respondents’ pupils as they interacted with the stimuli. Before the SP stage of the experiment began, participants were asked to take part in warm up and calibration exercises, to make sure that they were comfortable with the technology and that the eye-tracker was recording accurately and not being much of a distraction. The calibration involved a nine-point task, where participants followed a circle around the screen with their gaze (this is a standard approach in eye-tracking experiments).

Before testing began, participants were given a brief description of the task; however they were not shown a sample scenario image. For most of the respondents, the task was their first time encountering such a layout of information.

---

<table>
<thead>
<tr>
<th>Travel Time</th>
<th>Time +/-</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5 Hours</td>
<td>30 Minutes</td>
<td>€45</td>
</tr>
<tr>
<td>4 Hours</td>
<td>15 Minutes</td>
<td>€25</td>
</tr>
<tr>
<td>2.5 Hours</td>
<td>20 Minutes</td>
<td>€45</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Travel Time</th>
<th>Time +/-</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Car</td>
<td>2.5 Hours</td>
<td>30 Minutes</td>
</tr>
<tr>
<td>Coach</td>
<td>4 Hours</td>
<td>15 Minutes</td>
</tr>
<tr>
<td>Rail</td>
<td>2.5 Hours</td>
<td>20 Minutes</td>
</tr>
</tbody>
</table>

**FIGURE 1: STATED PREFERENCE SCENARIOS**
Sample
The sample used for this experiment was recruited mainly from a University population (staff and students), and therefore cannot be claimed to be representative of the Irish population as a whole. However, as this research is primarily concerned with the application of a technique and its potential for researchers, rather than the specific results of a given study, this was not deemed to be a major issue. A sample of 24 participants with a mean age of 26 years old was used for the purposes of this demonstration experiment. The sample was small in size due to resource constraints; however this research was designed to highlight the role that eye-tracking methods can play in survey design.

Analysis Techniques
Most eye-tracking software work on the basis of fixations and saccades. Fixations are periods of time when the gaze is at rest over a certain part of the image, and for the analysis presented in this study, they are defined at the pre-set value of 50 milliseconds. Saccades refer to the transitions between these fixations. These metrics form the underlying statistics for the generation of typical outputs. Chief amongst the methods of analysis used in eye-tracking are heat maps, areas of interest (AOIs), and scan paths. For more information into the specifics of the construction of these outputs, the reader is referred to Holmqvist (26).

Heat Maps
Heat maps provide a visual representation of where the gaze of either a given participant or the aggregate of the sample has fallen, based upon a selected underlying variable. Heat maps have the advantage of being highly intuitive and often easy to interpret with no prior knowledge of the area.

Areas of Interest (AOIs)
An AOI basically refers to an individual component/area of the visual presentation. For example, each text label for an alternative (such as “car” and “coach” on the left side of Figure 1) or an attribute for an alternative (such as the “travel time” cell for the “car” alternative”) can be considered as an AOI. Following the collection of the data, the analyst has the ability to compare a number of metrics relating to each of the areas of interest,
including visit count (the number of times an AOI is visited), the percentage fixated (the percentage of participants that gazed on a given AOI), the time to first fixation (in milliseconds), and the total visit duration (in milliseconds). The choice of metric to use will be dependent both upon the designers’ experience, and the requirements related to their research question.

**Scan Paths**

Scan paths are patterns created by linking the location of fixations together based upon the temporal order in which they occurred, to create a point-to-point mapping of the movement of the participants’ gazes across the image under consideration. This can also be used to create insightful animations that provide intuitive insight in relation to the participants’ behaviour. It is also possible to analyse these paths in a more statistically meaningful manner (26). Scan paths are particularly interesting to examine comparative behaviour with regard to attributes and alternatives, as they can often display behaviours such as checking and re-checking information.

Each of the above means of analysis has both advantages and drawbacks, and it is up to the study designer and analysts to select how to best utilize them to meet their own research objectives.

**RESULTS**

**Heat Map Results**

The heat maps are presented in Figure 3. These were derived from an examination of relative fixation duration, for two of the scenarios under consideration. These outputs are provided both in the form of traditional heat maps (the lower colour panel in Figure 3), and also burn through heat maps (the upper panel in Figure 3). Burn through heat maps begin as a black image and provide transparency to outline the underlying image in areas that received greatest degrees of attention. This allows the researcher to view the information that is of importance to the participants, rather than having it obscured by the colour distribution.

From the results shown for these two scenarios, it can be seen that, on average, all the attributes can be considered to be attended to, as can the two alternative labeling formats provided. From a designer’s point of view, this is a positive sign as it would appear to indicate that individuals are engaging with all the information that the scenario provides.
AOI results
Table 1 provides information on visit count, the percentage fixated, the time to first fixation, and the total visit duration.

Visit count indicates to the analyst the number of times an individual returned to a given AOI, suggesting a re-checking of information. This measure is similar for both the image-based and text-based labeling (of the alternatives) approaches, with scores of 6.82 and 5.83, for the images and text respectively. The “percentage fixated” metric indicates the percentage of participants that registered at least on fixation in a given AOI, and in both cases this is quite high. Time to first fixation reveals how quickly the information within a given AOI captured the participants’ attention, and this is much lower in this case for the images over the text, as would be expected given their highly visual nature. Finally, total visit duration gives an indication of the overall time spent reviewing the information within an AOI, and, on this metric, the alternative labels appear to be quite similar.

TABLE 1: LABEL COMPARISONS

<table>
<thead>
<tr>
<th></th>
<th>Images</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visit Count</td>
<td>6.82</td>
<td>5.83</td>
</tr>
<tr>
<td>Percentage Fixated</td>
<td>100%</td>
<td>92.31%</td>
</tr>
<tr>
<td>Time to First Fixation</td>
<td>849ms</td>
<td>6445ms</td>
</tr>
<tr>
<td>Total Visit Duration</td>
<td>2898ms</td>
<td>3463ms</td>
</tr>
</tbody>
</table>

Total Visits and Choice Consensus
While the visit counts recorded for each of the alternatives (each alternative being considered as an individual AOI) and attributes (each attribute for an alternative being considered as an individual AOI) do not map well onto the choices of participants, as may be expected based on the complex evaluation task underlying the scenarios, the choice information is not without value when assessing the results arising from the eye-tracking analysis. The visits to the relevant AOIs change as participants move through the scenarios, but the values seen for individual AOIs appear relatively stable in comparison to each other. This would suggest that there is some underlying trend that may be causing participants to gather more information for all AOIs in some specific scenarios than in others. One possible explanation for the variance in visit counts across the scenarios may be that, in some scenarios, the preferred choice for a given individual may not have been readily apparent, and therefore respondents recheck the attributes to help inform their decisions.
Table 2 outlines the percentage of participants who selected a given option for each of the choice scenarios. The choice scenarios presented in Table 2 represent the combination of the cost, trip and travel time variability. Using this data, it is possible to calculate a “non-consensus” variable for the alternative selections. By examining the percentage of respondents that selected a given attribute in each choice set and then creating a standard deviation based on these values and inverting them (dividing them into 1), it is possible to create a metric that represents the level of non-consensus between respondents for each scenario (Non-Con in Table 2). For example, in Scenario 1 there is a wide distribution of respondents across the three modes with values of 15.38%, 38.46%, and 46.15%, for car, bus, and rail respectively. An inverted standard deviation of these figures produces an output of 0.062, while, in Scenario 3, where all respondents selected the rail option, a value of 0.017 is produced. Thus, the more respondents agreed on the preferred choice for a given scenario, the lower is the value of the “non-con” variable, and the less the respondents agreed (the more evenly disturbed they are across their modal choices) the higher is the value of the “non-con” variable. It should be noted that the absolute values of this variable are less important than how the values for each scenario compare relative to each other. This variable essentially allows researchers to compare the levels of respondent agreement for a given scenario with the visit counts for the respective alternatives, in a manner that can be easily visually interpreted. By using the appropriate scaling factors, it is possible to map the change in these values with regard to each scenario, alongside the total visits for both the attributes and alternative values.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Car</th>
<th>Bus</th>
<th>Rail</th>
<th>STDEV</th>
<th>Non-Con</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1</td>
<td>15.38%</td>
<td>38.46%</td>
<td>46.16%</td>
<td>16.01</td>
<td>0.062</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>11.54%</td>
<td>84.62%</td>
<td>3.84%</td>
<td>44.58</td>
<td>0.022</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>0.00%</td>
<td>0.00%</td>
<td>100.00%</td>
<td>57.74</td>
<td>0.017</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>19.23%</td>
<td>0.00%</td>
<td>80.77%</td>
<td>42.19</td>
<td>0.024</td>
</tr>
<tr>
<td>Scenario 5</td>
<td>38.46%</td>
<td>57.69%</td>
<td>3.85%</td>
<td>27.28</td>
<td>0.037</td>
</tr>
<tr>
<td>Scenario 6</td>
<td>7.69%</td>
<td>84.62%</td>
<td>7.69%</td>
<td>44.42</td>
<td>0.023</td>
</tr>
<tr>
<td>Scenario 7</td>
<td>26.92%</td>
<td>57.69%</td>
<td>15.39%</td>
<td>21.87</td>
<td>0.046</td>
</tr>
<tr>
<td>Scenario 8</td>
<td>0.00%</td>
<td>17.39%</td>
<td>82.61%</td>
<td>43.55</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Figure 4 presents the change in total number of visits for the three alternatives under consideration as well showing the scaled non-consensus variable for each of the scenarios, as just discussed. This figure clearly displays how changes in levels of non-consensus track very well the number of times participants visited the information provided for each of the alternatives. This suggests a relationship between scenarios where respondents showed increased levels of non-consensus in their stated choices, and increased levels of attention being paid to the information provided. This may indicate that where the preferred choice is less obvious, more consideration of the data is required, and this is reflected in the number of visits to the relevant AOIs.
Scan Paths
Scan paths are a means of visually displaying the movement of the participants’ gaze across the visual presentation, which in our case represent the SP scenarios. Figure 5 provides a sample scan path for a single scenario and a single participant. The number of circles provides a sense of the number of fixations linked together, while the numbers represent the order in which they occurred. The size of the circle is related to the fixation duration. While scan paths are not particularly useful tools for presenting aggregate data for an entire sample, they can provide very useful insights into the information acquisition behaviours related to a given survey. The scan path displayed in Figure 5 represents one of the simplest in the dataset, in terms of the number of fixations recorded; however, it does clearly display the presence of both “within-alternative” and “within-attribute” data acquisition processes, and that the participant is switching between these strategies. For example, the specific individual in question starts with a fixation on cost, then moves to coach travel time, and so on. The horizontal gaze movement of the fixations is obvious (“within alternative” fixation), as is the vertical gaze movement of fixations (“within attribute” fixations).
DISCUSSION
The results presented in this paper highlight the ability of eye-tracking experiments to assess the information acquisition behaviours that occur when individuals are participating in SP experiments.

The analysis in this study provides researchers with a number of insights. First, this research demonstrates the potential analysis techniques that eye-tracking technology makes available to survey designers and researchers. The heat maps, both in the traditional and burn through formats, give researchers the ability to quickly scan the data to assess if there are any elements that are clearly not being attended to. These maps also allow researchers to easily understand which elements are receiving the most amount of attention. The AOIs provide the ability to gain more in-depth statistics about the elements that make up the images. Second, an examination of the results arising from the two sets of scenarios (textual versus image presentation of the alternatives) shows a clear attraction of the gaze toward images, rather than text. This suggests that an image-based presentation of alternatives may make the evaluation task easier and more efficient for respondents. Third, this study outlined how eye-trackers can be used to highlight relationships that exist between the nature of the scenarios, in this case levels of respondent consensus with regard to alternative selection, and the information assimilation behaviours of the participants. The results presented in Figure 4 indicate that, when the choice situation becomes more and more trickier (as measured by an increasingly even distribution of individual choices across the many alternatives), there is much more attention being paid to the AOIs under consideration. This would appear to suggest that respondents are checking and rechecking information when asked to make more difficult decisions. Inversely, these results suggest that, when one alternative is clearly the preferred option, the number of visits to the relevant AOIs decreases as less contemplation of options is required. Finally, the results provided by scan paths serve to highlight the complex comparison information that underlies such experiments. These paths show how respondents view information multiple times, returning to compare it with other elements of the scenario.

CONCLUSIONS AND RECOMMENDATIONS TO RESEARCHERS
This paper is designed to provide researchers with an insight into the role that eye-tracking can play in survey design, and specifically in the area of SP scenario design. The results and methods of analysis presented in this paper highlight some of the ways in which eye-tracking technology can be utilized to provide researchers with greater levels of information regarding how respondents interact with their scenarios, and assimilate the information that they strive to convey. While increased sample sizes may have the ability to produce more representative
results, even small samples can highlight important issues in scenario design, such as reduced levels of attendance on a given alternative or attribute.

The results and methods presented in this paper are intended to be an aid to researchers wishing to incorporate this new technology to improving their survey design; however they only highlight some of the possible applications of this analysis technique. The design of individual studies and the choice of methods of analysis are in the end the prerogative of the individual researcher, and certainly dictated by the needs relating to their research questions.

Survey design and piloting are essential elements of the data collection process. If these steps are not undertaken in a thorough manner, there is the possibility of unforeseen biases and other problems that cannot be rectified post data collection. Eye-tracking technology provides a new means of ensuring that surveys are able to collect information in a format appropriate to the research being undertaken. While eye-tracking is a highly effective means of gaining insight into how individuals interact with a given survey, it need not be undertaken in isolation; rather it can be considered to be another “instrument” in the designer’s toolbox, which can complement existing designing and piloting techniques. Survey design is not an exact science, and to some degree will always be based upon judgement calls. But we hope that the adoption of new techniques, such as eye-tracking, will contribute to improved design of survey collection instruments, and lead to better models to inform decision-making.

REFERENCES

8. Brazil, W., Caulfield, B., Does green make a difference? The potential role of smart phone technology in transport behaviour, Transportation Research Part C: Emerging Technologies, 37, 2013, p93 – 101
tracking and change detection in a case study with probiotic milk. LWT-Food Science and Technology, 68, 160-167


25. Eye Tribe Available at: https://theeyetribe.com/