Transport Reviews

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ttrv20

Utilising physiological data for augmenting travel choice models: methodological frameworks and directions of future research

Thomas O. Hancock & Charisma F. Choudhury

To cite this article: Thomas O. Hancock & Charisma F. Choudhury (2023): Utilising physiological data for augmenting travel choice models: methodological frameworks and directions of future research, Transport Reviews, DOI: 10.1080/01441647.2023.2175274

To link to this article: https://doi.org/10.1080/01441647.2023.2175274

Published online: 13 Feb 2023.

TRANSPORT REVIEWS
https://doi.org/10.1080/01441647.2023.2175274

Utilising physiological data for augmenting travel choice models: methodological frameworks and directions of future research

Thomas O. Hancock and Charisma F. Choudhury

Choice Modelling Centre and Institute for Transport Studies, University of Leeds, Leeds, United Kingdom

ABSTRACT
Recent technological and methodological advances have led to the possibility of a wider range of data being incorporated into travel choice models. In particular, physiological data such as eye-tracking information, skin conductance, heart rate recordings and electroencephalogram (EEG) have emerged as promising sources of information that could be used to gain insights into the decision-making process as well as the decision-maker's state of mind. However, research on methodologies to utilise these data sources and to integrate them with mobility data for advancing state-of-the-art travel behaviour models is still very limited. In this paper, we discuss the key benefits of using these emerging sources of physiological data, review applications of different types of physiological data and highlight their strengths and weaknesses. Particular attention is paid to two different generic frameworks for integrating these types of data into econometric choice models of travel behaviour. The first framework involves using physiological sensor data as indicators of latent variables while in the second framework, they are used as exogenous variables. We identify the research gaps and outline the directions for future methodological and applied research required to better utilise the physiological data for travel choice models.

ARTICLE HISTORY
Received 17 December 2021
Accepted 18 January 2023

KEYWORDS
Physiological sensor data; travel behaviour models; choice modelling; eye-tracking; stress measurement data; electroencephalogram (EEG)

1. Introduction

The travel decisions of individuals are affected by the attributes of the alternatives and the characteristics of the decision-maker. These characteristics range from socio-demographic factors (e.g. gender, age, income, etc.), attitudes (e.g. views about sustainable options), and traits (e.g. risk-taking propensity) to dynamic factors like the state of mind (e.g. stress level, excitement, etc.). Travel behaviour research to date has mainly focused on accounting for the effects of static characteristics on travel behaviour, with dynamic variations in decisions primarily captured by randomly distributed error terms representing intra-respondent heterogeneity (Ben-Akiva et al., 2007; Hess & Rose, 2009). Some

CONTACT Thomas O. Hancock T.O.Hancock@leeds.ac.uk Choice Modelling Centre and Institute for Transport Studies, University of Leeds, 34-40 University Road, Leeds LS2 9JT, United Kingdom

exceptions include work on modelling the relationship between travel decisions and mood and happiness (e.g. Le & Carrel, 2019), where retrospectively reported data on mood have been used as the dependent variable. However, like all other modelling work reliant on retrospectively reported and/or ratings data from participants, such models are prone to bias due to potential errors in the reporting and subjectiveness of the ratings. Further bias can also arise from the use of "point" (discontinuous) reports of mood since they may have strong reference dependence.

On the other hand, recent technological advances have led to newer sources of data for travel behaviour modelling which are passively generated and can consist of longer panels of repeated observations of the same decision-maker. In particular, physiological data such as skin conductance, heart rate recordings, etc. have emerged as promising sources of information to gain insights on the decision-making process as well as the decision-maker's condition or state of mind. Hence, these datasets open up further possibilities in incorporating inter and intra-respondent heterogeneity. However, thus far, there have been limited uses of complementary physiological data to aid the ability of a model to represent and capture travel behaviour.

Physiological data sources can be classified into two groups. The first group provides information regarding how an individual may process the world, such as eye-tracking information and brain activity recordings (e.g. electroencephalogram (EEG) or functional magnetic resonance imaging (fMRI)). The second group can provide indications about the decision-maker's psychological condition or state, such as heart rate, skin conductance, blood pressure, blink rate, facial expressions, etc. In addition to holding the promise of improving the models, the latter can provide insights regarding the well-being of the travellers.

The benefits of incorporating physiological sensor data in travel choice models are applicable in the contexts of both stated preference (SP) data and revealed preference (RP) data. Within the context of SP research, where the respondents are presented with hypothetical scenarios and asked to make choices in controlled settings, additional physiological sensor data can be used specifically to better understand the thought process a decision-maker goes through between being presented with information and then making a choice. Further, within the context of SP research, where the respondents are presented with hypothetical scenarios and asked to make choices in controlled settings, the responses may be subject to "warm glow" effect (Nunes & Schokkaert, 2003) and/or respondents may try to state "correct" choices and attempt to hide the "actual" ones they would make in real-world settings. Physiological data on the other hand is not subject to obfuscation, i.e. a decision-maker cannot "hide" physiological indicators. For example, Millen and Hancock (2019) found that participants could not effectively alter their eye fixation patterns to pretend they did not recognise familiar faces, and measures such as heart-rate and skin conductance are similarly autonomic and not easily possible to consciously control. Hence, collecting and analysing physiological sensor data in conjunction with SP responses can help the analyst to identify which choices demonstrably show preferences, and which are more strongly influenced by emotions, stress or other factors that may cause an individual to choose an alternative they may not usually choose. For example, Hancock et al. (2022) demonstrate that more stressed drivers make more "random" choices. Physiological sensor data can also motivate respondents to have more careful engagement, as would the use of cheap talk or oaths (Cummings &

Taylor, 1999; Crastes dit Sourd et al., 2018). For example, Mahieu et al. (2016) used lie detectors to assess the truthfulness of participants' responses and Cherchi et al. (2020) used EEG readings to understand when decision-makers found choice tasks easy or difficult. Further, EEG has also been used to interpret a decision-maker's confidence in their choices in perceptual tasks (Krumpe et al., 2018). This has the additional benefit of avoiding the requirement of advanced methodological frameworks to circumvent endogeneity issues in the joint modelling of choice and choice certainty ratings (Habib, 2017). Another clear benefit of using physiological sensor data is that it can be linked with underlying decision-making processes. For example, eye-tracking information gives insights on information processing strategies (Ryan et al., 2018) and fMRI and EEG recordings can help infer learning (Frank et al., 2015) and consumer decision-making (Golnar-Nik et al., 2019) processes. This allows for improving behavioural insights and may also allow for a better understanding of different possible decision-making strategies across individuals in the controlled settings of SP.

In the context of RP data, where it is harder for modellers to capture the full range of external factors that may impact choices made in the real-world, physiological data could be used to better account for heterogeneous behaviour at the level of the individual decision-maker, for example by capturing the stress levels of a road-user travelling through busy traffic to understand their route choice behaviour. For example, Xu et al. (2018) demonstrate that eye-tracking information could be used to assess how tired a driver is in real-time, which can be used to better explain their route and lane choice, gap-acceptance decisions, etc.

Finally, the respective limitations of the RP and SP data have motivated travel behaviour researchers to combine RP and SP data to complement each other (see Ben-Akiva & Morikawa, 1990; Buckell & Hess, 2019; Lizana et al., 2021 for details), though their uptake beyond academia is still limited (Batley et al., 2019). One alternative to combining SP and RP is to again consider physiological data (which when collected anonymously may have less privacy issues attached to them compared to detailed socio-demographic information) and use it as a measure of cross-validation of preferences, thus potentially removing some concerns regarding the realism of behaviour in SP settings. For example, Brunyé and Gardony (2017) demonstrate that eye-tracking information such as pupil diameter can be used to predict choice certainty. This can be a potential way to link choices made in RP and SP settings.

Furthermore, both the traditional RP and SP data have limited information about the "softer" factors that affect an individual's travel behaviour such as mental states like stress, anger or frustration. Evidence from psychology suggests that these have a significant impact on choices (Garfinkel et al., 2016; Starcke et al., 2012). For example, controlled experiments have demonstrated that stress increases risk-taking propensity and reduces variety-seeking behaviour of individuals (Buckert et al., 2014) and numerous studies have demonstrated the impact of stress or distraction on driving behaviour (e.g. Paschalidis et al., 2018). However, so far there has only been limited efforts in collecting such data in the context of generic travel behaviour modelling – primarily through retrospective reporting (Carrel et al., 2016) and experience sampling (Carrel et al., 2017), both of which are prone to substantial reporting and measurement errors.

Given the clear promise outlined above from using physiological data for understanding behaviour, it is unsurprising that applications using this kind of data are becoming increasingly common. However, it is less clear that travel behaviour modellers, relative to modellers from other disciplines, are making as much use of physiological data. Furthermore, as there is a wide variety of types of physiological data and also a number of theories regarding the impact certain measures should have on behaviour, it is understandable that there is also a wide variety of methods for incorporating this data into models. This is consequently one of the aims of this paper: to review results from previous behavioural modelling with physiological data and consider different possible frameworks for including such data. This paper also outlines some important steps for future travel behaviour modelling involving physiological data, including a discussion on some barriers that have thus far limited the use of this data, not least of which is the lack of an established framework for its incorporation into standard choice models.

The rest of this paper is arranged as follows. Firstly, we outline two general frameworks for the inclusion of physiological data in choice models. Next, we discuss different types of physiological data in turn, detailing previous applications and frameworks for their integration into choice models, relating these different methods to our generalised frameworks. We then discuss opportunities and next steps for using physiological data within travel behaviour models, before drawing some conclusions.

2. Two generalised frameworks for the inclusion of physiological data in choice models

There exists a wide variety of possible methods to analyse the relationship between physiological sensor data and human behaviour. However, in the domain of choice modelling (both travel and beyond), the models developed to date can be broadly categorised as falling under one of the two generalised frameworks. Both of these generalised frameworks include a number of different functions (arrows in Figures 1 and 2) to represent

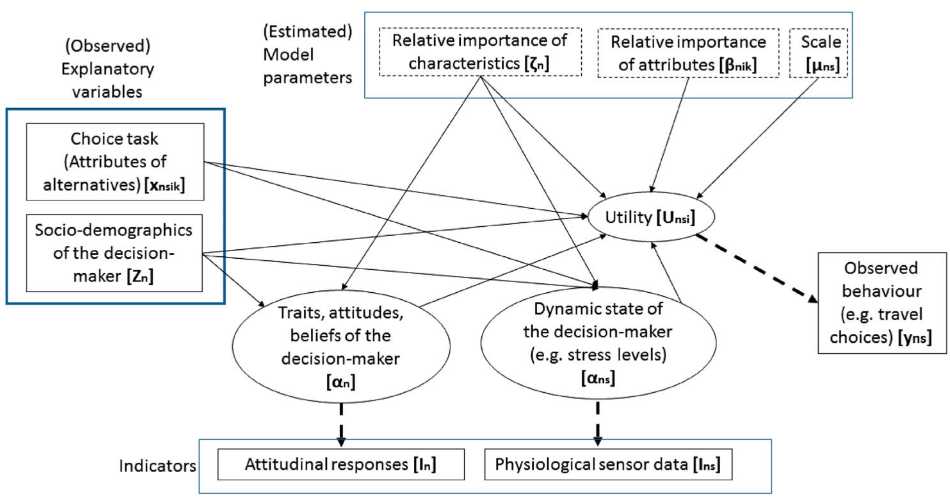

Figure 1. Framework 1: an integrated choice and latent variable model with physiological sensor data as indicators. The indices used are for individual, *n*, alternative, *i*, attribute, *k*, and choice task, *s*.

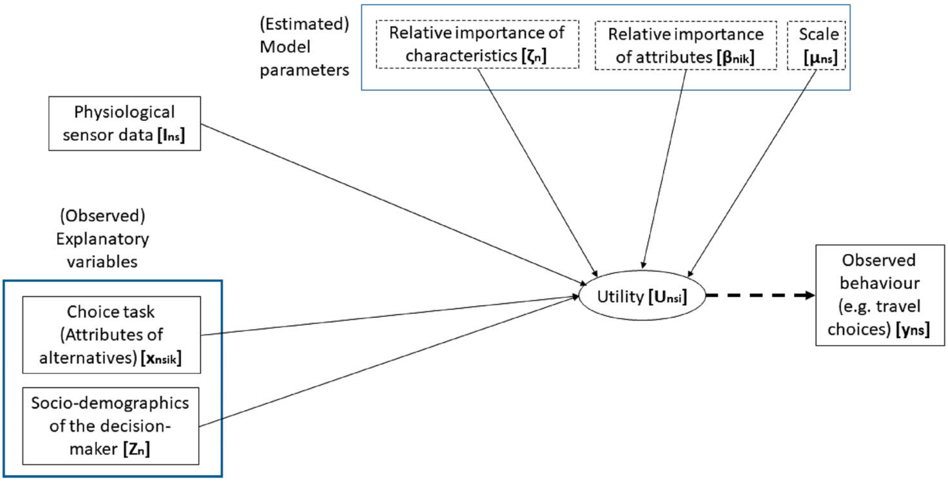

Figure 2. Framework 2: a process and choice model with physiological sensor data used as direct explanators.

the full set of possible elements of decision-making when specifying models incorporating physiological sensor data. A summary of these functions is given in Table 1. It may be noted that in reality, due to the unavailability of the data or identification issues, most applications thus far have included some subset of the functions included in these generalised frameworks (as detailed in Section 3). In both of our frameworks, it is assumed that some number of observed variables (in rectangles in the figures) and estimated parameters (in dashed rectangles) are specified together in some functional form to calculate the utilities of different alternatives. Measurement equations linking the latent variables (in ovals) with the observed outputs are represented by dashed arrows (note that associated measurement errors are excluded in the figures for the sake of simplified presentation).

The key difference between the two discussed frameworks is the placement and use of physiological sensor data. Depending on the application context and problem of interest, one framework may be preferable compared to the other.

Table 1. A summary of the functions (arrows) visualised in Figures 1 and 2, and the key differences between the frameworks.

Framework 1	Framework 2
Functional relationships	
(A) $U_{n,s,i} = f(\alpha_{n,s}, Z_n, x_{n,s,i,k}, \zeta_n, \beta_{n,i,k}, \mu_{n,s})$ (C) $y_{n,s} = f(U_{n,s,i})$ (D) $I_{n,s} = f(\alpha_{n,s})$ (E) $\alpha_{n,s} = f(Z_n, \zeta_n, x_{n,s,i,k})$	(B) $U_{n,s,i} = f(I_{n,s}, Z_n, x_{n,s,i,k}, \zeta_n, \beta_{n,i,k}, \mu_{n,s})$ (C) $y_{n,s} = f(U_{n,s,i})$
Key advantages	
Allows for measurement error in physiological sensor data Can capture correlations between physiological sensor and choice data	Simpler model to estimate Potential to have better explanatory power of choice model
Key disadvantages	
May have limited explanatory variables to define latent variables α	Predictions using the model require simulation of physiological sensor data

In the first framework, physiological data is used in a latent variable approach and is assumed to be an indicator. Explanatory variables are thus used to optimise the likelihood of observing both the choice outcomes and the physiological indicators. This framework is consequently equivalent to standard Integrated Choice and Latent Variable (ICLV) models (Ben-Akiva et al., 2002; Vij & Walker, 2016) but with physiological indicators replacing or in addition to standard indicators such as attitudinal responses. This framework is visually represented in Figure 1. The figure illustrates possible functional links between known (solid rectangles), estimated (dashed rectangles) and latent variables (solid ovals). Here it is assumed that physiological data is to be treated as indicators within an ICLV framework. The latent variable(s) are used to predict the physiological indicators alongside the "traditional" indicators (such as attitudinal responses) and the observed choices. The latent variables used in typical ICLV models are assumed to be static across choice contexts (i.e. not dependent on the choice task, s) and only depend on the characteristics of the decision-maker, n. As a contrast, the latent variables that are used to explain physiological sensor data are "dynamic". These latent variables can thus be informed by both characteristics of the decision-maker and variables related to the choice context (e.g. task difficulty). They can consequently be used to help explain the decision-making process. For example, they could represent the perception of information (a decision-maker's perceived risk, Bogacz et al., 2021) or the factors impacting the processing of information (stress levels, Paschalidis et al., 2019).

In the second framework, physiological data is instead incorporated as an explanatory variable. This data thus helps inform the utility of alternatives, sometimes indirectly through impacting the value of the estimated parameter(s) through some specified function. Thus, for example, eye-tracking information may directly inform attribute attendance and thus effectively capture individual-specific relative importances of different attributes (Pike et al., 2020).

This framework is visually represented in Figure 2.

It should be noted that attitudinal responses are not included in Framework 2 as it has been widely acknowledged that in order to avoid the measurement errors and endogeneity biases associated with attitudinal responses, it is crucial to use ICLV frameworks (Ashok et al., 2002; Ben-Akiva et al., 2002). Though the measurement errors are also potentially a problem for physiological sensor data (Krucien et al., 2017), these are not controllable by the decision-maker themselves (e.g. Millen & Hancock, 2019). Hence they are less likely to lead to endogeneity issues and subsequent modelling issues.

Further, for both frameworks, a modeller would not likely use all components detailed in the figures. However, should the analyst have extensive data on, for example, a wide range of decision-maker characteristics, the identification of more model parameters and the addition of further latent constructs may be possible.

We now give a mathematical summary of the graphical representations of the frameworks given in the figures in Table 1 where the subscripts n, s, i, k denote decision-maker, scenario, alternative and attribute respectively. In this table, we give generalised functions (arrows in the figures) that have been used to incorporate physiological data. We denote typical indices for each variable, which may change across contexts. For example, physiological sensor data such as heart rate would only vary across individuals and the scenario ($I_{n,s}$) but eye-tracking data may also include information related to an attribute (i.e. how much attention it receives) thus we would instead have $I_{n,s,k}$.

It is well worth noting that a majority of applications utilising physiological sensor data use some combination of the functions in Table 1, with a subset of the elements in each function depending on what information is available to the modeller. For example, a typical mode choice model may define the utility for alternatives based on characteristics of a decision-maker (i.e. their age, gender, etc), the attributes of the alternatives, and estimated parameters for the relative importance of these attributes and characteristics, meaning that we have $U_{n,s,i} = f(x_{n,s,i,k}, Z_n, \beta_{n,i,k}, \zeta_n)$.

Table 1 also gives key advantages and disadvantages of each framework.

As the first framework also aims to predict physiological sensor data, it requires methods for defining the latent variables. Within standard implementations of ICLV models for understanding attitudinal responses, these variables are often set as a function of the characteristics of the decision-maker. However, the construction of these latent variables can be trickier in models predicting physiological sensor data. For example, Krucien et al. (2017) used a latent variable to represent the perception of information, allowing the model to capture the correlation between how much a decision-maker looked at the information and the importance of the information. However, this latent variable was based on an error term only in the final model. One possible alternative is to use scenario attributes to determine the latent variable. For example, Paschalidis et al. (2019) used driving variables (e.g. relative speeds) to define a latent variable that represented how stressed an individual was, that in turn influenced stress measurements and also choices.

The key change for the second framework is that observed physiological sensor data is instead used as explanatory variables within utility functions. The main advantage here is that there is the possibility of the model having better choice explanatory power (e.g. Hancock et al., 2022), Whilst the models are also easier to estimate, this approach causes issues when the model is to be used in a predictive context. This is a result of possibly complex or rudimentary assumptions being required for the generation of simulated physiological sensor data to inform the new choice contexts. As a contrast, ICLV models can be used for forecasting choice behaviour without the requirement of indicator data for the new hypothetical scenarios (Vij & Walker, 2016).

In the subsequent section, we consider a number of different types of physiological sensor data, detailing how they have been included in choice models and how these applications relate to the generalised frameworks presented above.

3. Physiological sensor data and their implementations within choice models

We now review the precise methods for how physiological data can be incorporated into choice models, giving specific examples of the generalised equations given in Table 1. We consider distinct types of physiological sensor data separately. For each type of data, we first provide an overview of the type of data where we summarise the sub-types of data and how they have been used in the literature. We discuss some general findings from using the data, followed by a particular focus on how data has been mathematically incorporated within choice models and relating past approaches to the two generalised frameworks defined in Section 2.

3.1. Eye-tracking data

3.1.1. Overview

Interest and research involving tracking eye movement has existed for over a hundred years, with psychologists initially studying eye movement to understand the reading process (e.g. Huey, 1898). However, it was not until the 1970s that eye-tracking research came to prominence, through the domain of experimental psychology, when researchers started attempting to link perception and mental processes (Płużyczka, 2018). In the context of decision-making, eye-tracking information has often been used to understand how decisions are made by considering how decision-makers visually process information (Chen et al., 2021; Krajbich et al., 2012). This also allows for the validation (or the lack thereof) of competing theories of decision-making behaviour, with eye-tracking information frequently used to argue for or against various assumptions within choice models (e.g. Mullett & Stewart, 2016).

In Table 2, example applications of eye-tracking data are given to demonstrate the variety of sub-types of data/revealed information, to detail example applications following our two model frameworks, and to illustrate the breadth of uses of eye-tracking data. It is notable that the type of eye-tracking information that a researcher may want to use varies depending on the application context (and often subject to the type of available eye-tracker). A further key question facing a researcher wanting to make use of eye-tracking information is *how* to empirically include it within their behavioural model. In previous research, both variants of the generic framework presented in Section 2 have been used: as a direct variable (i.e. Figure 2) and as an indicator of a latent variable (i.e. Figure 1). The former is more prevalent in non-econometric modelling literature. On the other hand, in econometric choice models, it has been acknowledged that the eye-tracking data can have measurement errors and can consequently need to be treated as an indicator (Krucien et al., 2017). The majority of applications for including eye-tracking information can thus be summarised using a subset of the functions are given in Table 1 (demonstrated in the following subsections).

In each case in Table 2, we specify whether the modelling work (if combined with choice data) falls into Framework 1 (Figure 1 or Framework 2 (Figure 2). We discuss each of these applications in detail in the relevant subsequent subsection.

3.1.2. Subtypes of eye-tracking data

Given the huge number of studies investigating visual attention (multiple reviews of eye-tracking research exist across different disciplines, for example, Ziv, 2016), it is unsurprising that the type of eye-tracking information used varies significantly. This includes, but is not limited to fixation location, fixation length, blink rate and duration and pupil size measurements. In addition to the available equipment, the type of collected data is largely impacted by what is feasible in the experimental setting. For example, in a dynamic or virtual setting, it may not be possible to easily work out exactly what the decision-maker is looking at and thus less specific information such as the number of times a decision-maker blinks or the general gaze direction (left, right, etc.) may

Table 2. A summary across key work utilising eye-tracking data, considering types of data, revealed information, how this has been used in mathematical modelling and the relation to our identified methodological frameworks.

Manuscript	Data type	Revealed information	Data & Application	Framework	Equations	Mathematical modelling/statistics
Ballco et al. (2019)	Fixation location/ length	Detailed focus of decision-maker	SP, food preferences	2	B, C	Probabilities of choice alternatives based on visual attention
Benedetto et al. (2011)	Blink rate, duration and pupil size measurements	Concentration or fatigue levels	Driving simulator, driver performance	n/a	n/a	Repeated measures ANOVA to test differences across tasks
Chavez et al. (2018)	Fixation location/ length	Detailed focus of decision-maker	SP, food preferences	2	B, C	Probabilities of choice alternatives based on visual attention
Chen et al. (2021)	Gaze direction	Focus of decision-maker	RP, shopping behaviour	n/a	D	Prediction of visual attention based on shop layout, product popularity, etc.
Dudinskaya et al. (2020)	Fixation location/ length	Detailed focus of decision-maker	SP, food preferences	2	B, C	Probabilities of choice alternatives based on visual attention
Fisher (2017)	Fixation location/ length	Detailed focus of decision-maker	SP, food preferences	2	B, C	Probabilities of choice alternatives based on visual attention
Hancock et al. (2022)	Gaze direction	Focus of decision-maker	Driving simulator, gap acceptance decisions	2	B, C	Probabilities of choice alternatives based on visual attention
Krajbich et al. (2012)	Fixation location/ length	Detailed focus of decision-maker	SP, purchasing behaviour	2	B, C	Probabilities of choice alternatives based on visual attention
Krucien et al. (2017)	Fixation location/ length	Detailed focus of decision-maker	SP, health preferences	1	A, C, D, E	ICLV framework predicting both choices and attribute fixation times
Li et al. (2018)	Gaze concentration (based on gaze direction)	Concentration or fatigue levels	Driving simulator, driver performance	n/a	D	Regressions for micro-steering activity based on gaze concentration
Merat et al. (2012)	Blink rate and duration	Concentration or fatigue levels	Driving simulator, driver performance	n/a	n/a	Repeated measures ANOVA to test differences across tasks
Mullett and Stewart (2016)	Gaze direction	Focus of decision-maker	Simulated data, binary choice tasks	n/a	D	Simulation of gaze direction (and choice response time) patterns under different modelling assumptions
Palinko et al. (2010)	Pupil size measurements	Concentration or fatigue levels	Driving simulator, driver performance	n/a	n/a	Repeated measures ANOVA to test differences across tasks
Pike et al. (2020)	Fixation location/ length	Detailed focus of decision-maker	SP, vacation preferences	2	B, C	Probabilities of choice alternatives based on visual attention
Shimojo et al. (2003)	Gaze direction	Focus of decision-maker	SP, (human) face preferences	n/a	D	Regressions for gaze direction based on face characteristics
Spinks and Mortimer (2015)	Fixation location/ length	Detailed focus of decision-maker	SP, medical treatment preferences	n/a	D	Prediction of attribute non-attendance based on task complexity, sociodemographics, etc.
Uggeldahl et al. (2016)	Gaze shifts	Focus of decision-maker	SP, food preferences	2	B, C	Probabilities of choice alternatives based on gaze shifts

*Note that "n/a" in the Framework column refers to cases where methods other than choice modelling were used.

instead be used. For example, when a purchaser scans different items in a supermarket, Chen et al. (2021) found that items on the right drew more attention than those on the left, and that items on a shelf below eye-level rather than at eye-level received the most attention. Studies of driving behaviour are similarly dynamic, and thus blink rate is often used to assess a driver's concentration or performance level (Merat et al., 2012). This is particularly important in the context of investigating how distracted a driver is, with blink duration (Benedetto et al., 2011), pupil size measurements (Palinko et al., 2010) and gaze concentration (measured as a function of the standard deviation of the horizontal gaze position, Li et al., 2018) also used for this purpose, and Khan and Lee (2019) summarising techniques and applications of incorporating eye-movement data specifically for use within advanced driving assistance systems.

As a contrast, in a laboratory setting, where the decision-maker is facing a screen, precise information on *saccades*, numbers of fixations and fixation length may be recorded. As a result, a modeller may have precise information regarding how long a decision-maker looks at particular alternatives or particular attributes, thus having a proxy for the relative importance of different features. This has led to multiple studies incorporating eye-movement data with a view to understanding attribute non-attendance (Chavez et al., 2018; Spinks & Mortimer, 2015). Many studies also use eye-movement information to try and understand the decision-making process itself, with research in cognitive psychology finding a late onset bias (Shimojo et al., 2003), where decision-makers tend to look more at the alternative they are about to choose just before they choose it (also known as a gaze-cascade effect).

3.1.3. Models based on Framework 1

The models based on Framework 1 involve using eye-tracking variables as indicators of latent constructs within an ICLV framework. This makes it possible to calculate probabilities in prediction contexts without the need of generating any simulated eye-tracking data (as required in models based on Framework 2). This is, however, not a common approach within the domain of eye-tracking research, with one key example being that of Krucien et al. (2017), who found that preferences for "harder to process" attributes varied more significantly with changes in underlying visual attention. Under their framework:

$$U_{n,s,i} = f(x_{n,s,i,k}, \alpha_{n,k}, \beta_k \cdots) \tag{1a}$$

$$I_{n,k} = f(\alpha_{n,k}) \tag{1b}$$

where the latent variable $\alpha_{n,k}$ is dependent on an individual's processing strategy (though the final model by Krucien et al. (2017) represented this simply by noise) and is then used to calculate both the utility of an alternative and the total fixation time for an attribute k across choice tasks completed by participant n. Krucien et al. (2017) argue that this alternative approach leads to a number of benefits, including removing the assumption of a deterministic relationship between visual attention and individuals' preferences, with preferences instead impacted by the underlying level of information intake (thus recognising the fact that an individual may not necessarily be thinking about what they are looking at).

3.1.4. Models based on Framework 2

There are many more examples of applications where eye-tracking data is used to help predict choice data, in line with Framework 2. In particular, some of the first applications of eye-tracking data in the context of decisions research were performed within cognitive psychology. In general, models developed in this discipline are typically not based on econometric theory, meaning they do not necessarily incorporate concepts of utility (for a review of these models, see Busemeyer et al., 2019, and for a review on insights from eye-tracking research within cognitive psychology, see Krajbich, 2019). As a result, for some models, the quality of an alternative is simply related directly to the visual attention it receives:

$$U_{n,s,i} = f(D_{n,s,i}, F_{n,s,i}, \ \ldots), \tag{2}$$

where D and F are the relative shares of attention duration and fixations, respectively, an alternative i receives over the course of the decision-making process (thus relative importance/marginal utility parameters are not estimated nor included). This is the approach assumed by the visual attention model (see Equation 1, Chavez et al., 2018), where the probability of choosing an alternative is simply:

$$P_{n,s,i} = D_{n,s,i}. \tag{3}$$

This means that the probability of choosing a particular alternative is set directly as the proportion of time spent looking at the alternative in comparison to the time spent looking at all alternatives. The attentional drift-diffusion model (Krajbich et al., 2012) also assumes that the quality of an alternative is based only on the visual attention it receives, and has subsequently been used for predicting the probability of choosing an individual's preferred snack (from a set of images of different snacks, Fisher, 2017), where the attributes of the different alternatives may, in any case, be relatively difficult to directly compare.

The direct inclusion of eye-movement data can also be effective in the utility-based models that are typically used for modelling responses to stated preference (SP) surveys. For example, the utility of an alternative could be a function of the value of different attributes combined with the visual attention for the attributes. In studies on preferred tourism destinations (Pike et al., 2020) and food (yoghurt) preferences (Ballco et al., 2019), the utility under a mixed logit model was set as a sum of these components:

$$U_{n,s,i} = f(x_{n,s,i,k}, \beta_{n,i,k}, I_{n,s}, \ldots) \tag{4a}$$

$$= \sum_{k=1}^{K}\left(x_{n,s,i,k} \cdot \beta_k\right) + \sum_{k=1}^{K}\left(D_{n,s,k} \cdot x_{n,s,i,k} \cdot \gamma_k\right), \tag{4b}$$

where D is the total fixation duration for an attribute k in choice task s made by decision-maker n, $x_{n,s,i,k}$ gives the value for attribute k for alternative i, and β_k and γ_k are marginal utility and relative visual importance parameters, respectively, for attribute k, to be estimated. This framework is effective in that should the γ parameters have insignificant estimates, the model collapses to a standard mixed logit model without the incorporation of eye-tracking information.

This setup could also be used in the context of attribute non-attendance research, though an alternative is to include attributes in the utility calculation only if they receive a certain share of visual attention. Chavez et al. (2018) compared their visual attention model to a "conditional logit" model, where utility for an alternative was defined:

$$U_{n,s,i} = \sum_{k=1}^{k} \left(x_{n,s,i,k} \cdot \beta_k \cdot (1 - ANA_{n,s,k}) \right), \tag{5}$$

where $ANA_{n,s,k}$ is an indicator function for attribute non-attendance, set to a value of 1 if the relative share of visual attention that an attribute receives is less than 10% of the total decision-making time ($D_{n,s,k} < 0.1$). A similar approach was also adopted by Dudinskaya et al. (2020), who used an equivalent function but with $ANA_{n,s,k}$ set to a value of 1 if there are less than 2 fixations on attribute k.

A similar approach was used by Hancock et al. (2022), where it was demonstrated that drivers who look more to the left (towards oncoming traffic) than they usually do whilst driving assign a higher importance to the size of gaps when choosing which gap to accept at unsignalised intersections. In their model, possible measurement error was included by assuming that the percentage of "attention time" (AT) given to considering the size of the gap was only proportional rather than equivalent to the fixation time (D):

$$AT_{n,s,i,k} = f(\alpha, D_{n,s,i,k}), \tag{6}$$

where a high estimate for α corresponds to attention time matching fixation time, and a low estimate suggests attention is random, and not dependent on fixations. Positive, significant estimates were found for α, suggesting that individuals did indeed assign more importance to the attributes that they looked at.

Eye-tracking information has also been used within decision-making research as a means to understand choice certainty, with Brunyé and Gardony (2017) demonstrating that fixations, saccades, and pupil diameter could all be used as measures for uncertainty. In the context of a utility-based framework, choice certainty can be impacted through adjustments to the scale parameter. For example, Uggeldahl et al. (2016) set scale as a function of gaze shifting:

$$\mu_{n,s} = f(I_{n,s}, \ \ldots) \tag{7a}$$

$$= 1 + \alpha_{GS} \cdot GS_{n,s} + f(A_{n,s}) \tag{7b}$$

where $GS_{n,s}$ is a z-score transformation of the number of gaze shifts (movements between recorded fixations) and $A_{n,s}$ is a set of alternative choice set features including whether there is pictorial or text representation, choice set order, etc (described in full detail by Uggeldahl et al., 2016, Section 2.5). Under the above specification, a negative estimate for α_{GS} (as indeed found by Uggeldahl et al., 2016) implies that the more gaze shifts recorded within a choice task, the lower the scale within the model, corresponding to a "less certain" response.

3.2. Biomarker data

3.2.1. Overview

The most frequently used biomarker data includes heart rate, skin conductance, respiration, blood volume pulse, salivary cortisol and muscular activity. Applications of these data are relatively new in comparison to eye-tracking data and are mostly limited to inferring stress and fatigue levels and their impact on a decision-maker's behaviour (for a review, see Starcke & Brand, 2012). They have been primarily used in the context of driving safety research (Crawford, 1961). Given the highly consequential nature of poor driving, it is not surprising that the majority of studies within a transport context that aim to measure how stressed an individual is, focus on relating stress with driving behaviour. A particular interest has been the development of passive methods to detect how stressed a driver is. Physiological sensor indicators offer such an opportunity, particularly as technological advances have resulted in increasingly non-intrusive sensors, and alternative measures such as travel diaries are subject to reporting biases and rely on individuals actually remembering what has happened (Gulian et al., 1990). Before discussing some example implementations of stress measurement data in detail, we give example applications of the use of biomarker data in Table 3. We again aim to demonstrate the variety of subtypes of data/revealed information, to detail example applications following our two model frameworks, and to illustrate the breadth of uses of biomarker data.

3.2.2. Subtypes of biomarker data

Multiple physiological sensors have been utilised when the aim is to understand how stressed a driver is, with Healey and Picard (2005) providing one of the first advanced case studies of real-world driving behaviour. In their experiment, electrocardiogram, electromyogram, skin conductance and respiration were recorded for 24 drivers as they drove through a set route in the greater Boston area. Healey and Picard (2005) demonstrated that these measures were highly correlated with a stress metric based on observed driving conditions, road features and other behaviours of the driver, such as bumps in the road, head turns, turns in the road and having to stop the car. Chen et al. (2017) further explored this data through the use of more advanced methods, including the use of machine learning models to more accurately predict when a driver was stressed or not. Outside of travel behaviour research, similar methods have also been used to assess the mental state and stress levels of patients with anxiety disorders (Katsis et al., 2011b), with Smets et al. (2018) providing a detailed review of some of the challenges faced when trying to measure stress levels outside of laboratory settings in general. In particular, they highlighted the importance of accounting for individual-level variations in sensor readings and discussed the difficulties of finding 'base levels' as a result of the impact of physical activity on these measures.

Stress measurement data has also been implemented in the context of research on decision-making. For example, Nowacki et al. (2019) demonstrated that men made more risky choices in comparison to women after being subject to a 'cold pressor' test, which was designed to deliberately induce stress, by raising blood pressure, heart rate and salivary cortisol. However, heart-rate was also higher for participants who made less risky lottery choices (Fooken & Schaffner, 2016), implying that those who do not feel stressed are more likely to make risky choices as they possibly do not worry as much about their choices. Stress levels have also been demonstrated to have an effect in moral choice scenarios. For example, higher stress results in less utilitarian choices in moral dilemmas (Starcke et al., 2012) and individuals with lower resting heart rates were found to give lower ratings of anticipated guilt should they commit crimes in confrontational situations as well as giving lower predicted probabilities of being convicted (Armstrong & Boutwell, 2012).

Table 3. A summary across key work utilising biomarker data, considering types of data, revealed information, how this has been used in mathematical modelling and the relation to our identified methodological frameworks.

Manuscript	Data type	Revealed information	Data & Application	Framework	Equations	Mathematical modelling/statistics
Armstrong and Boutwell (2012)	Resting heart rate	Emotional involvement	Survey, emotion/action ratings	n/a	n/a	Regressions for prediction of ratings based on resting heart rate
Chen et al. (2017)	Muscular activity, breathing rate and or depth	Stress levels	RP, driving stress levels	n/a	D	Correlation between stress levels and driving conditions
Fooken and Schaffner (2016)	Heart rate variability	Emotional involvement	SP, risky decision-making	2	B, C	Probabilities of choice alternatives based on heart rate
Hancock et al. (2022)	Electrocardiogram (heart rate); Skin conductance (sweat rate)	Stress levels	Driving simulator, gap acceptance decisions	2	B, C	Probabilities of choice alternatives based on stress levels
Healey and Picard (2005)	Muscular activity, breathing rate and or depth	Stress levels	RP, driving stress levels	n/a	D	Correlation between stress levels and driving conditions
Katsis et al. (2011b)	Blood volume pulse, heart rate	Mental state	RP, emotions during therapeutic sessions	n/a	n/a	Comparison of classification of mental state (expert psychologist vs machine learning models)
Nowacki et al. (2019)	Salivary cortisol, blood pressure, heart rate	Stress levels	SP, risky decision-making	n/a	n/a	Repeated measures ANOVA to test differences across tasks
Paschalidis et al. (2018)	Heart rate, sweat rate	Stress levels	Driving simulator, gap acceptance decisions	2	B, C	Probabilities of choice alternatives based on stress levels
Paschalidis et al. (2019)	Heart rate, sweat rate	Stress levels	Driving simulator, car-following behaviour	1	A, C, D, E	ICLV framework predicting both choices and stress levels
Starcke et al. (2012)	Heart rate	Stress levels	SP, moral decision-making	n/a	n/a	Repeated measures ANOVA to test differences across groups
Tarabay and Abou-Zeid (2021)	Heart rate	Stress levels	Driving simulator, red-light violations	1	A, C, D, E	ICLV framework predicting both choices and stress levels

*Note that “n/a” in the Framework column refers to cases where methods other than choice modelling were used.

3.2.3. Models based on Framework 1

Latent variable approaches in-line with Framework 1 (Figure 1) have also been used for the incorporation of stress indicator data. Paschalidis et al. (2019) treated stress as a latent unobserved variable in their car-following behavioural model, thus demonstrating that estimated stress levels could explain heart rate, blood volume pulse and skin conductance recordings ($I_{n,s}$) through measurement equations:

$$I_{n,s} = f(\alpha_{n,s}), \tag{8a}$$

$$= f(ZA_n, x_{n,s,k}) \tag{8b}$$

where the latent variable ($\alpha_{n,s}$) is based on a function of sociodemographics (ZA_n) and driving variables ($x_{n,s,k}$). The choice to accelerate/decelerate was then also a function of latent stress:

$$U_{n,s,i} = f(\alpha_{n,s}, x_{n,s,k}). \tag{9}$$

Tarabay and Abou-Zeid (2021) utilised a similar framework, showing that driving performance measures, as well as physiological indicators, could be explained by latent variables representing a driver's level of stress.

3.2.4. Models based on Framework 2

Whilst most uses of stress indicator data have been either to classify stress levels or within simple statistical analyses of decision-making, there have been a few incorporations of stress indicator data within discrete choice models. Paschalidis et al. (2018) demonstrated that heart rate and skin conductance measures could be used to show that drivers under stress were more likely to accept a gap when crossing an unsignalised intersection. They incorporated these measures within the utility to accept a gap (error terms are omitted):

$$U_{n,s} = \beta_k \cdot x_{n,s,k} + \zeta \cdot ZA_n + \theta \cdot I_{n,s}, \tag{10}$$

where $x_{n,s,k}$ is a set of gap-specific variables (e.g. size of the gap), ζ and θ are the estimated coefficients for the impact of sociodemographics, ZA_n, and the physiological variables, $I_{n,s}$. A number of physiological sensor indicators were tested, with the participant's normalised heart rate and skin conductance responses found to be significant (higher levels resulted in higher utilities for accepting the gap). Hancock et al. (2022) built upon this work by demonstrating that these indicators could also be used within a decision field theory (DFT) model (Busemeyer & Townsend, 1993; Hancock et al., 2021). The size of the normal error term (which has a similar effect to utility scale, thus is also labelled μ) within the DFT model was adjusted depending on the amount of stress:

$$\mu_{n,s} = f(I_{n,s}), \tag{11}$$

with results demonstrating that higher levels of stress resulted in "more random" choice behaviour under the model.

3.3. EEG and fMRI data

3.3.1. Overview

Whilst psychologists have been at the forefront of research utilising brain imaging techniques, these methods have gradually been used in the context of decision research and then also travel behaviour research. The two major types, electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI), differ in terms of their spatial and temporal resolutions. fMRI has a high spatial but low temporal resolution, whilst EEG has a lower spatial but higher temporal resolution.

In Table 4, example applications are given to demonstrate the variety of sub-types of data/revealed information, to detail example applications following our two model frameworks, and to illustrate the breadth of uses of EEG/fMRI data.

3.3.2. Subtypes of neural data

fMRI has predominantly been used to identify which parts of the brain contribute to the completion of different tasks (Rodriguez et al., 2015; Zysset et al., 2006). This includes many studies on decision-making processes, with the dorsomedial prefrontal cortex playing a key role in risky decision-making (Rao et al., 2011), the ventromedial prefrontal cortex and ventral striatum reflecting value, value comparison and confidence (De Martino et al., 2013; Gluth et al., 2015) and responses in the amygdala and orbitofrontal cortex correlating with ambiguity in choices (Hsu et al., 2005). Though fMRI has a lower temporal resolution than EEG, its outputs can be compared across different decisions. Rodriguez et al. (2015) demonstrated that larger differences in gambling options lead to less activation from fMRI outputs (implying an easier decision) and Gluth et al. (2015) demonstrated that stated food ratings correlated with fMRI outputs.

Meanwhile, within the context of travel behaviour research, specific choice task difficulty has been tested with EEG recordings, with Cherchi et al. (2020) demonstrating that EEG could be used to infer the difficulty of SP tasks on car purchase decisions. EEG has also been used to investigate how participants detect collision threats (Markkula et al., 2021) and can be used to predict when a driver will perform an emergency brake during simulated driving (Haufe et al., 2011). Furthermore, it has been integrated into neural networks and Markov models, respectively, to detect fatigue in drivers (Karuppusamy & Kang, 2020) and pilots (Wu et al., 2021).

3.3.3. Models based on Framework 1

In the most advanced travel behaviour application of EEG data thus far, Bogacz et al. (2021) incorporate EEG responses within a hybrid choice model to understand cycling decisions in a virtual reality experiment. The (latent) relative risk felt by the individual as they are cycling is then used to approximate alpha-wave activity recordings from the EEG data. Their model thus has similar components to that of Paschalidis et al. (2019), with alpha activity ($I_{n,s}$) estimated through measurement equations:

$$I_{n,s} = f(\alpha_{n,s}), \tag{12a}$$

$$= f(x_{n,s,k}) \tag{12b}$$

where the latent (risk) variable ($\alpha_{n,s}$) is based on a function of driving variables ($x_{n,s,k}$) such as the distance to the next junction and whether there is a car near the cyclist. The utility

Table 4. A summary across key work utilising neural data, considering types of data, revealed information, how this has been used in mathematical modelling and the relation to our identified methodological frameworks.

Manuscript	Data type	Revealed information	Data & Application	Framework	Equations	Mathematical modelling/statistics
Bogacz et al. (2021)	EEG: alpha wave activity	Implied risk preferences	Cycling simulator, cycling speed decisions	1	A, C, D, E	ICLV framework predicting both choices and relative risk
Cherchi et al. (2020)	EEG: activity from many locations	Implied effort in decision-making	SP, car preferences	n/a	n/a	Repeated measures ANOVA to test differences across tasks
De Martino et al. (2013)	fMRI: ventromedial prefrontal cortex	Implied confidence in choice	SP, food preferences	n/a	D	Correlation between neural activity and stated choice confidence ratings
Gluth et al. (2015)	fMRI: activity from many locations	Location of neural activity associated with specific tasks	SP, food preferences	n/a	D	Regressions for neural activity based on (independent) choice modelling outputs
Haufe et al. (2011)	EEG: activity from many locations	Timing of decision-making	Driving simulator, emergency brake responses	n/a	n/a	Comparison of neural activity time and action (breaking) times
Hsu et al. (2005)	fMRI: activity from many locations	Location of neural activity associated with specific tasks	SP, risky and ambiguous decision-making	n/a	D	Regressions for neural activity based on choice task attributes
Karuppusamy and Kang (2020)	EEG: alpha wave activity	Implied drowsiness of driver	Driving simulator, drowsiness when driving	n/a	n/a	Prediction of driver drowsiness based on EEG activity
Lusk et al. (2016)	fMRI: ventromedial prefrontal cortex activity	Implied preference for alternatives	SP, food preferences	2	B, C	Probabilities of choice alternatives based on fMRI activity
Markkula et al. (2021)	EEG: activity from many locations	Timing of detection of stimuli	Perceptual, collision threat detection	n/a	D	Prediction of neural activity times based on task variables
Rao et al. (2011)	fMRI: dorsomedial prefrontal cortex activity	Location of neural activity associated with specific tasks	SP, risky decision-making	n/a	D	Regressions for neural activity based on choice task attributes
Rodriguez et al. (2015)	fMRI: activity from many locations	Location of neural activity associated with specific tasks	SP, intertemporal choice	n/a	D	Regressions for neural activity based on (independent) choice modelling outputs
Telpaz et al. (2015)	EEG: theta wave activity	Implied preference for alternatives	SP, consumer preferences	n/a	D	Correlation between neural activity and preferences
Webb et al. (2013)	fMRI: medial prefrontal cortex activity	Implied preference for alternatives	SP, consumer preferences	2	B, C	Probabilities of choice alternatives based on fMRI activity
Webb et al. (2019)	fMRI: medial prefrontal cortex activity	Implied preference for alternatives	SP, consumer preferences	2	B, C	Probabilities of choice alternatives based on fMRI activity
Wu et al. (2021)	EEG: activity from many locations	Implied drowsiness of pilot	Flight simulator, measurement of fatigue	n/a	n/a	Prediction of pilot drowsiness based on EEG activity

*Note that 'n/a' in the Framework column refers to cases where methods other than choice modelling were used.

(at time point s) for decisions regarding cycling speed (to brake, maintain speed, wait or accelerate) is then also a function of latent risk:

$$U_{n,s,i} = f(\alpha_{n,s}, x_{n,s,i,k}). \tag{13}$$

3.3.4. Models based on Framework 2

Lusk et al. (2016) developed a random utility model that incorporated fMRI outputs. First, they measured the participant's ventromedial prefrontal cortex response to different attributes of boxes of eggs, presented individually (low prices, high prices, free range or caged hens). Second, participants completed a number of stated choice tasks, where their preference for different alternatives could be related to their fMRI responses. Thus, the utility for a box of eggs was defined:

$$U_{n,s,i} = f(x_{n,s,i,k}, \beta_k, I_{n,k,l}), \tag{14a}$$

$$= (\beta_{cost1} + \beta_{cost2} \cdot I_{n,cost,l}) \cdot cost_{n,s,i} + (\beta_{fr1} + \beta_{fr2} \cdot I_{n,fr,l}) \cdot free_{n,s,i} + \delta_i, \tag{14b}$$

where $cost_{n,s,i}$ is the cost of alternative i in choice task s faced by respondent n, $free_{n,s,i}$ is a dummy variable set to a value of 1 if the eggs are from free-range hens, and 0 otherwise, δ_i is an alternative specific constant and $I_{n,cost,l}$ and $I_{n,fr,l}$ are the ventromedial prefrontal cortex responses to different attributes levels, l (from the first phase of the experiment). This specification, as with Equation (4) for eye-tracking information, is thus based on a standard utility specification but with additional parameters for capturing the relative importance of attributes based on physiological information.

Furthermore, a "neural random utility model" (NRUM, Webb et al., 2013) was developed to predict which of a pair of products a consumer will choose on the basis of differences in fMRI readings. In these models (which have typically been based on probit models), the utility for a single item is simply based on the fMRI output ($I_{n,s,i}$):

$$U_{n,s,i} = I_{n,s,i} \cdot \mu, \tag{15}$$

where the only estimated term is the scale. Extensions to these models include incorporating measurement error (Webb et al., 2019). Similar approaches have also been adopted with the use of EEG data, which as a contrast to fMRI, has low spatial resolution but high temporal resolution. It is consequently easier to use for studying phenomena such as preference formation. EEG is thus also often used within neuromarketing research (e.g. Telpaz et al., 2015).

4. Next steps

In this section, we first summarise the key barriers that have thus far hindered the use of physiological sensor data within travel choice models. We then consider how we might overcome these barriers and discuss a number of opportunities presented by the further use of this kind of data.

4.1. Barriers for implementation

Whilst the previous section has demonstrated a number of successful uses of physiological sensor data, there are a number of barriers that have thus far stopped the widespread use of this kind of data in the context of mathematical modelling of travel choices. These barriers can be grouped into three key areas of concern: (1) interpretation of physiological sensor data; (2) methodological; (3) economic.

Firstly, it can be difficult to relate physiological sensor data directly to specific observed events or effects, even in the context of laboratory settings. This is due to the fact that physiological sensors produce noisy signals which are typically very large due to the high resolution. This is particularly an issue for brain imaging data such as EEG, which can easily produce thousands of observations per second, thus requiring extensive data processing (Sanders et al., 2020; Sanei & Chambers, 2013), and is subject to perturbations due to voluntary (e.g. moving limbs), non-voluntary actions (e.g. blinking eye) and environmental changes (e.g. change in lighting levels). Consequently, to relate data to observable effects, significant restraining or simplification of choice tasks is often required. This becomes a particular challenge in the context of real-world scenarios, where the travel contexts are less controllable and often only partially observable. Though the level of noise is comparatively less in the case of biomarker data, they are often more susceptible to external environmental factors. For example, the effect of weather and humidity can pose risks in getting consistent skin conductance data and appropriate corrections factors may be warranted. Further, such data have high levels of state-dependence with spikes triggered by earlier events leaving residuals that affect the subsequent measurements (i.e. an elevated heart rate caused by a stressful event can take a while to return to normal). This leads to significant complications in terms of direct interpretation of the outputs from physiological sensors. Moreover, the properties of the biomarkers tend to be person-specific. For example, the same level of stress can lead to different amplitudes of heart rate variations among different people. To address this issue, person-specific base levels of each biomarker need to be pre-determined (as done for heart rate in Paschalidis et al., 2019) which results in the requirement of large sample sizes at the individual-level before these variables can have significant effects in models (as highlighted by Smets et al., 2018). There are also large differences in brain activity between hypothetical and real-world contexts (Camerer & Mobbs, 2017). The wider variety of stimuli in the real-world contexts could not only impact measurements but be impossible to account for. For example, Engström et al. (2017) show that increased cognitive load (e.g. from a conversation on the phone) can lead to arousal of heart-rate and result in worse driving decisions. A further complication for use of these types of data in real-world settings is that synchrony of equipment is crucial, as some events of interest may have temporary effects that are observable in physiological sensor data for only a few milliseconds (this is particularly the case for perceptual cues, see Jamal et al., 2015) and can be missed.

Secondly, incorporating the physiological sensor data in choice models is complex and has several methodological challenges. In fact, this is partly the reason why in a major share of the previous studies (both in travel and other choice contexts), physiological sensor data have been used to test only simple statistical measures such as how much outputs from these sensors correlate with other observed variables. For the behavioural models that do incorporate physiological sensor data, it is far simpler to use them as explanatory variables, but more challenging to include them as indicators. On the contrary, theoretical arguments and empirical evidence (i.e. Krucien et al., 2017) imply that among the two approaches (directly observable vs. latent), the latent approach should be adopted. This difficulty arises from the problem of finding alternative explanatory variables to use to explain or inform estimates for physiological indicators. For example, there are no a priori socio-demographic variables that can easily be linked to where different individuals will look in a given choice context. As highlighted by the different equations in the previous section, methodological issues also exist in the form of a lack of an established practice for the incorporation of physiological sensor data, both in terms of which type of framework to adopt, and the functional form in which the sensor data is included. The existing studies adopt many different approaches, with the result that analysts looking to utilise sensor data have too many options for possible functional forms to incorporate the data. This is partly a result of the fact that utility-based choice models have few clear conceptually obvious options for where sensor data should fit into the model. This is even the case for eye-tracking information, for which different levels of attention could impact attributes/alternatives in different ways, both psychologically (conceptually) and statistically (the mathematical impact of parameters) within a model. This issue is of particular concern when considering brain-imaging data, where having thousands of data points per second means that mapping a signal to some cognitive process can itself be a complex process, before even introducing this data into a choice model.

Finally, experimental costs are of course a clear restriction on the use of physiological sensor data. As well as upfront monetary costs of equipment, there are heavy labour costs associated with the use of physiological sensor data. Much time is required to set up participants with the sensors, which may also be uncomfortable or intrusive (resulting frequently in limited participant numbers), and for some types of sensors, the knowledge requirement for interpreting data is inhibiting (Gramfort et al., 2013). For example, some knowledge of physics, signal processing, statistics, and numerical methods is required to interpret EEG data.

Overall, whilst there may seem many barriers, the following section highlights methods for mitigating issues, as well as alternative opportunities presented by the further utilisation of physiological sensor data.

4.2. Opportunities

The barriers listed in the previous sub-section can be used to formulate the agenda for future research for overcoming the barriers. The key research opportunities in the field are mapped onto the challenges presented in Figure 3 and are detailed below.

Firstly, with regards to issues pertaining to how to incorporate physiological data, there is clear scope for further empirical comparisons of our respective frameworks (visualised in Figures 1 and 2), with Krucien et al. (2017)'s results favouring the adoption of a latent variable framework. Results may differ across contexts and depending on the type of physiological sensor data, though hybrid approaches appear to be the more flexible option given they allow for measurement errors in the sensor data. It should also be noted that whilst there has thus far been limited uptake of latent variable approaches for physiological sensor data, this is not a result of it being difficult to predict this data on its own. There are many examples in Tables 2–4 where the focus is on the prediction of physiological sensor data alone (without corresponding choice data). This demonstrates that there is substantial scope for future development of models following Framework 1.

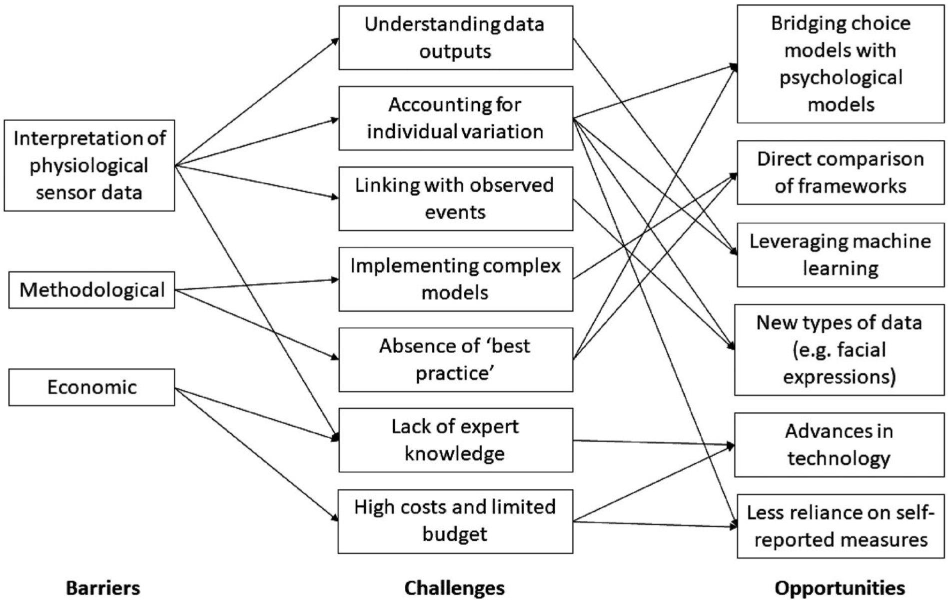

Figure 3. Overview of the key barriers, challenges and possible opportunities for the integration of physiological sensor data into choice models.

Further modelling framework tests are also possible through implementations of physiological sensor data into psychological choice models (Hancock et al., 2022). Within the context of cognitive psychology, a number of psychological choice models exist that have "process parameters" (Busemeyer et al., 2019; Hancock et al., 2021) that may provide clearer conceptual links with physiological sensor data. For example, decision field theory has attribute attention weights as well as attribute importance weights (Hancock et al., 2021) in addition to a parameter that could be a proxy for how long the decision-maker considers their alternatives (Hancock et al., 2019), meaning that elements of the decision-making process have their own distinct parameters. Thus further use of physiological sensor data will allow for further testing of the behavioural assumptions of psychological choice models. This may lead to a wider gap between the psychological choice models and those built on econometric theory compared to the small differences that are found while using choice-only data (Hancock et al., 2021).

Future work could also incorporate both of the above elements: testing latent variable versions of psychological choice models (which, as far as the authors are aware, have never been tested). These models, for example, could then have a latent process driving 'speed of information processing', which would then impact the parameters within these models that are related to the length of time spent on the decision-making process (which in turn impacts the probabilities of choosing the different alternatives, e.g. the parameter for the number of preference updating steps in a decision field theory model, Hancock et al., 2019). Such models can be also extended to quantify well-being impacts by linking them with measurements of brain activity (e.g. EEG outputs). This may also help split the decision-making process into distinct parts: the perception of information and the processing of information. This is in line with research from neuroscience, where fMRI studies demonstrate that there are different regions of the brain associated with the integration of information and the interpretation of it (Zysset et al., 2006).

Additionally, with regard to reducing the noise in sensor data, it may be possible that supplementary information or more advanced modelling specifications could be used for the calibration of person-specific parts of the model. For example, Tarabay and Abou-Zeid (2021) use a dynamic model to account for serial correlation in their hybrid model quantifying the impact of stress on driving decisions, Wu et al. (2021) demonstrate that hidden Markov models can be used to dynamically account for the accumulation of stress over time, and Gao et al. (2019) demonstrate that machine learning methods can be used to relate EEG measures to how fatigued a driver is. Further development of advanced methods such as these may allow for easier interpretations of the dynamic effects observed in physiological sensor data.

There are also a number of other sources/types of physiological data that have thus far received very little attention within the context of travel behaviour research. For example, facial expressions could be used, with Katsis et al. (2011a) demonstrating that the emotional states of drivers could be established with a wearable system that could record facial expressions, which can then be classified with machine learning methods. Furthermore, facial electromyography (facial EMG), which records face muscular movements, has been used to interpret a decision-maker's reaction in consumer stated choice tasks (Rasch et al., 2015). Alternatively, voice data could be used, with machine learning methods now advanced enough to, for example, infer stress levels from the pitch and word intervals (Adams et al., 2014). These new data sources may help to capture heterogeneity in behaviour within individuals, as well as helping to more clearly link other physiological sensor data with specific observed events.

Furthermore, research involving self-reported measures (e.g. well-being (Abou-Zeid & Ben-Akiva, 2012; Carrel et al., 2016) or personality types (Boyce et al., 2019; Calastri et al., 2017)), are typically subject to significant reporting biases. Physiological sensor data have significant promise to complement or replace these measures with potentially more dependable data. For example, EEG can produce biomarkers of well-being (Chilver et al., 2020) that can be used to recognise emotion by the deployment of supervised machine learning techniques (Liu et al., 2011). Similarly, preliminary research suggests that it may be possible to use EEG to identify some elements of an individual's personality (Li et al., 2020). Thus, combined with physiological measures of stress, a researcher may be able to directly estimate participants' state of mind.

Finally, continuous advances in technology may reduce the monetary costs of equipment, or alternatively provide more advanced equipment that is more user-friendly and easier to implement (e.g. less intrusive eye-tracking equipment could be used to monitor real-world driving behaviour (Khan & Lee, 2019)).

5. Conclusions

In this paper, we start by discussing a number of important benefits for travel behaviour analysis that will arise from the use of physiological sensor data in mathematical models of travel choices. We then give two generalised frameworks for the inclusion of this data in choice models, with the key difference between the frameworks being the placement of the sensor data within the model. There are key advantages to using the data exogenously to explain choices, such as the fact that they can result in improved prediction of choices (e.g. Hancock et al., (2022) demonstrate that eye-tracking and stress measurement data can help predict a driver's gap acceptance decisions). However, the use of latent variables avoids potential measurement error (a decision-maker may not be considering the attributes of an alternative that they are looking at) though comes at a cost of the requirement of additional data (usually further characteristics of a decision-maker) to explain the latent variable that drives both choice data and sensor measurement data.

We discuss a number of results observed from the incorporation of physiological sensor data, demonstrating the wide variety of methodological setups that have been used in previous research, and relating these to either one of the generalised frameworks presented in this paper. Notably, most applications thus far fall into the category of using physiological sensor data exogenously (Framework 2) though there has been a recent shift towards the use of latent variables. The aims of the analyst likely influence the choice of framework, as, for example, additional data or variables (e.g. choice task difficulty) may be required to also explain sensor data in a latent variable approach. However, forecasting of choices using a model designed in line with Framework 2 would require some possibly ad-hoc modelling assumptions as well as intensive simulations of sensor data to generate choices for new choice scenarios.

We conclude by detailing a number of barriers inhibiting the use of physiological data and giving some potential solutions, as well as further opportunities for future research. Whilst the work in this paper focuses on how to integrate physiological sensor data into choice models, it should be noted that a separate study on how to use this data more generally for travel behaviour analysis would likely help increase the use of physiological sensor data for travel behaviour research.

Overall, key avenues for future research include further direct comparisons of methodologies, enriching choice modelling with concepts of neuropsychology and mathematical psychology, the use of alternative types of physiological sensor data not yet integrated into choice models (e.g. facial expressions) and further utilisation of physiological sensor data to help explain both inter and intra-respondent heterogeneity in real-world contexts. These will undoubtedly advance our understanding of travel behaviour and result in models that are both more valid and more realistic.

Acknowledgement

The authors are grateful to four anonymous reviewers for their helpful feedback on previous versions of the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by UKRI Future Leader Fellowship [MRT020423/1].

ORCID

Thomas O. Hancock http://orcid.org/0000-0002-8922-0901
Charisma F. Choudhury http://orcid.org/0000-0002-8886-8976

References

Abou-Zeid, M., & Ben-Akiva, M. (2012). Well-being and activity-based models. *Transportation*, *39*(6), 1189–1207. https://doi.org/10.1007/s11116-012-9387-8

Adams, P., Rabbi, M., Rahman, T., Matthews, M., Voida, A., Gay, G., Choudhury, T. & Voida, S. (2014). Towards personal stress informatics: Comparing minimally invasive techniques for measuring daily stress in the wild. In *Proceedings of the 8th international conference on pervasive computing technologies for healthcare* (pp. 72–79).

Armstrong, T. A., & Boutwell, B. B. (2012). Low resting heart rate and rational choice: Integrating biological correlates of crime in criminological theories. *Journal of Criminal Justice*, *40*(1), 31–39. https://doi.org/10.1016/j.jcrimjus.2011.11.001

Ashok, K., Dillon, W. R., & Yuan, S. (2002). Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of Marketing Research*, *39*(1), 31–46. https://doi.org/10.1509/jmkr.39.1.31.18937

Ballco, P., de Magistris, T., & Caputo, V. (2019). Consumer preferences for nutritional claims: An exploration of attention and choice based on an eye-tracking choice experiment. *Food Research International*, *116*, 37–48. https://doi.org/10.1016/j.foodres.2018.12.031

Batley, R., Bates, J., Bliemer, M., Börjesson, M., Bourdon, J., Ojeda Cabral, M., Chintakayala, P.K., Choudhury, C., Daly, A., Dekker, T., Drivyla, E., Fowkes, T., Hess, S., Heywood, C., Johnson, D., Laird, J., Mackie, P., Parkin, J., Sanders, S., … Worsley, T. (2019). New appraisal values of travel time saving and reliability in Great Britain. *Transportation*, *46*(3), 583–621. https://doi.org/10.1007/s11116-017-9798-7

Ben-Akiva, M., Abou-Zeid, M., & Choudhury, C. (2007). Hybrid choice models: From static to dynamic. In *Triennial symposium on transportation analysis VI*. Citeseer.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D. S., & Daly, A. (2002). Hybrid choice models: Progress and challenges. *Marketing Letters*, *13*(3), 163–175. https://doi.org/10.1023/A:1020254301302

Ben-Akiva, M., & Morikawa, T. (1990). Estimation of switching models from revealed preferences and stated intentions. *Transportation Research Part A: General*, *24*(6), 485–495. https://doi.org/10.1016/0191-2607(90)90037-7

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*(3), 199–208. https://doi.org/10.1016/j.trf.2010.12.001

Bogacz, M., Hess, S., Calastri, C., Choudhury, C. F., Mushtaq, F., Awais, M., Nazemi, M., Van Eggermond, M. A. B., & Erath, A. (2021). *Modelling risk perception using a dynamic hybrid choice model and brain-imaging data: Application to virtual reality cycling*. Working paper, Institute for Transport Studies, University of Leeds.

Boyce, C., Czajkowski, M., & Hanley, N. (2019). Personality and economic choices. *Journal of Environmental Economics and Management*, *94*, 82–100. https://doi.org/10.1016/j.jeem.2018.12.004

Brunyé, T. T., & Gardony, A. L. (2017). Eye tracking measures of uncertainty during perceptual decision making. *International Journal of Psychophysiology*, *120*, 60–68. https://doi.org/10.1016/j.ijpsycho.2017.07.008

Buckell, J., & Hess, S. (2019). Stubbing out hypothetical bias: Improving tobacco market predictions by combining stated and revealed preference data. *Journal of Health Economics*, *65*, 93–102. https://doi.org/10.1016/j.jhealeco.2019.03.011

Buckert, M., Schwieren, C., Kudielka, B. M., & Fiebach, C. J. (2014). Acute stress affects risk taking but not ambiguity aversion. *Frontiers in Neuroscience*, *8*, 82. https://doi.org/10.3389/fnins.2014.00082

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in Cognitive Sciences*, *23*(3), 251–263. https://doi.org/10.1016/j.tics.2018.12.003

Busemeyer, J. R., & J. T. Townsend (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432. https://doi.org/10.1037/0033-295X.100.3.432

Calastri, C., Hess, S., Carrasco, J. A., Choudhury, C., & Daly, A. (2017). Maintaining social contacts over time: The role of life-course events, personality traits and tie strength. In *International choice modelling conference 2017*.

Camerer, C., & Mobbs, D. (2017). Differences in behavior and brain activity during hypothetical and real choices. *Trends in Cognitive Sciences*, *21*(1), 46–56. https://doi.org/10.1016/j.tics.2016.11.001

Carrel, A., Mishalani, R. G., Sengupta, R., & Walker, J. L. (2016). In pursuit of the happy transit rider: Dissecting satisfaction using daily surveys and tracking data. *Journal of Intelligent Transportation Systems*, *20*(4), 345–362. https://doi.org/10.1080/15472450.2016.1149699

Carrel, A., Sengupta, R., & Walker, J. L. (2017). The San Francisco travel quality study: Tracking trials and tribulations of a transit taker. *Transportation*, *44*(4), 643–679. https://doi.org/10.1007/s11116-016-9732-4

Chavez, D., Palma, M., & Collart, A. (2018). Using eye-tracking to model attribute non-attendance in choice experiments. *Applied Economics Letters*, *25*(19), 1355–1359. https://doi.org/10.1080/13504851.2017.1420879

Chen, L.-l., Zhao, Y., Ye, P.-F., Zhang, J., & Zou, J.-Z. (2017). Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers. *Expert Systems with Applications*, *85*, 279–291. https://doi.org/10.1016/j.eswa.2017.01.040

Chen, M., Burke, R. R., Hui, S. K., & Leykin, A. (2021). Understanding lateral and vertical biases in consumer attention: An in-store ambulatory eye-tracking study. *Journal of Marketing Research*, *58*(6), 1120–1141. https://doi.org/10.1177/0022243721998375

Cherchi, E., Vuong, Q. C., & Stergiou, A. (2020). Using EEG to understand how our brain elaborate information in stated choice experiments: Easy versus hard tasks in the choice of vehicles. bioRxiv.

Chilver, M. R., Keller, A. S., Park, H. R., Jamshidi, J., Montalto, A., Schofield, P. R., Clark, C. R., Harmon-Jones, E., Williams, L. M., & Gatt, J. M. (2020). Electroencephalography profiles as a biomarker of wellbeing: A twin study. *Journal of Psychiatric Research*, *126*, 114–121. https://doi.org/10.1016/j.jpsychires.2020.04.010

Crastes dit Sourd, R., Zawojska, E., Mahieu, P.-A., & Louviere, J. (2018). Mitigating strategic misrepresentation of values in open-ended stated preference surveys by using negative reinforcement. *Journal of Choice Modelling*, *28*, 153–166. https://doi.org/10.1016/j.jocm.2018.06.001

Crawford, A. (1961). Fatigue and driving. *Ergonomics*, *4*(2), 143–154. https://doi.org/10.1080/00140136108930515

Cummings, R. G., & Taylor, L. O. (1999). Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. *American Economic Review*, *89*(3), 649–665. https://doi.org/10.1257/aer.89.3.649

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–110. https://doi.org/10.1038/nn.3279

Dudinskaya, E. C., Naspetti, S., & Zanoli, R. (2020). Using eye-tracking as an aid to design on-screen choice experiments. *Journal of Choice Modelling*, *36*, 100232. https://doi.org/10.1016/j.jocm.2020.100232

Engström, J., Markkula, G., Victor, T., & Merat, N. (2017). Effects of cognitive load on driving performance: The cognitive control hypothesis. *Human Factors*, *59*(5), 734–764. https://doi.org/10.1177/0018720817690639

Fisher, G. (2017). An attentional drift diffusion model over binary-attribute choice. *Cognition*, *168*, 34–45. https://doi.org/10.1016/j.cognition.2017.06.007

Fooken, J., & Schaffner, M. (2016). The role of psychological and physiological factors in decision making under risk and in a dilemma. *Frontiers in Behavioral Neuroscience*, *10*, 2. https://doi.org/10.3389/fnbeh.2016.00002

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. *Journal of Neuroscience*, *35*(2), 485–494. https://doi.org/10.1523/JNEUROSCI.2036-14.2015

Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., & Zuo, S. (2019). EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(9), 2755–2763. https://doi.org/10.1109/TNNLS.5962385

Garfinkel, S. N., Zorab, E., Navaratnam, N., Engels, M., Mallorquí-Bagué, N., Minati, L., Dowell, N. G., Brosschot, J. F., Thayer, J. F., & Critchley, H. D. (2016). Anger in brain and body: The neural and physiological perturbation of decision-making by emotion. *Social Cognitive and Affective Neuroscience*, *11*(1), 150–158. https://doi.org/10.1093/scan/nsv099

Gluth, S., Sommer, T., Rieskamp, J., & Büchel, C. (2015). Effective connectivity between hippocampus and ventromedial prefrontal cortex controls preferential choices from memory. *Neuron*, *86*(4), 1078–1090. https://doi.org/10.1016/j.neuron.2015.04.023

Golnar-Nik, P., Farashi, S., & Safari, M.-S. (2019). The application of EEG power for the prediction and interpretation of consumer decision-making: A neuromarketing study. *Physiology & Behavior*, *207*, 90–98. https://doi.org/10.1016/j.physbeh.2019.04.025

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267. https://doi.org/10.3389/fnins.2013.00267

Gulian, E., Glendon, A., Matthews, G., Davies, D., & Debney, L. (1990). The stress of driving: A diary study. *Work & Stress*, *4*(1), 7–16. https://doi.org/10.1080/02678379008256960

Habib, K. N. (2017). Improving choice model parameter estimates by jointly modelling the sp choices with corresponding elicited certainty ratings. *Transportation Research Part A: Policy and Practice*, *95*, 305–319. https://doi.org/10.1016/j.trb.2016.07.013

Hancock, T. O., Choudhury, C. F., & Hess, S. (2022). *Secret in their eyes: Incorporating eye-tracking and stress indicator data into choice models.*. Working paper, Institute for Transport Studies, University of Leeds.

Hancock, T. O., Hess, S., & Choudhury, C. F. (2019). A careful respondent or an uncertain response: Disentangling confounding sources of increased deliberation time using decision field theory. Working paper, Institute for Transport Studies, University of Leeds.

Hancock, T. O., Hess, S., Marley, A., & Choudhury, C. F. (2021). An accumulation of preference: Two alternative dynamic models for understanding transport choices. *Transportation Research Part B: Methodological*, *149*, 250–282. https://doi.org/10.1016/j.trb.2021.04.001

Haufe, S., Treder, M. S., Gugler, M. F., Sagebaum, M., Curio, G., & Blankertz, B. (2011). EEG potentials predict upcoming emergency brakings during simulated driving. *Journal of Neural Engineering*, *8*(5), 056001. https://doi.org/10.1088/1741-2560/8/5/056001

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, *6*(2), 156–166. https://doi.org/10.1109/TITS.2005.848368

Hess, S., & Rose, J. M. (2009). Allowing for intra-respondent variations in coefficients estimated on repeated choice data. *Transportation Research Part B: Methodological*, *43*(6), 708–719. https://doi.org/10.1016/j.trb.2009.01.007

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, *310*(5754), 1680–1683. https://doi.org/10.1126/science.1115327

Huey, E. B. (1898). Preliminary experiments in the physiology and psychology of reading. *The American Journal of Psychology*, *9*(4), 575–586. https://doi.org/10.2307/1412192

Jamal, W., Das, S., Maharatna, K., Pan, I., & Kuyucu, D. (2015). Brain connectivity analysis from EEG signals using stable phase-synchronized states during face perception tasks. *Physica A: Statistical Mechanics and Its Applications*, *434*, 273–295. https://doi.org/10.1016/j.physa.2015.03.087

Karuppusamy, N. S., & Kang, B.-Y. (2020). Multimodal system to detect driver fatigue using EEG, gyroscope, and image processing. *IEEE Access*, *8*, 129645–129667. https://doi.org/10.1109/Access.6287639

Katsis, C., Goletsis, Y., Rigas, G., & Fotiadis, D. (2011a). A wearable system for the affective monitoring of car racing drivers during simulated conditions. *Transportation Research Part C: Emerging Technologies*, *19*(3), 541–551. https://doi.org/10.1016/j.trc.2010.09.004

Katsis, C. D., Katertsidis, N. S., & Fotiadis, D. I. (2011b). An integrated system based on physiological signals for the assessment of affective states in patients with anxiety disorders. *Biomedical Signal Processing and Control*, *6*(3), 261–268. https://doi.org/10.1016/j.bspc.2010.12.001

Khan, M. Q., & Lee, S. (2019). Gaze and eye tracking: Techniques and applications in adas. *Sensors*, *19* (24), 5540. https://doi.org/10.3390/s19245540

Krajbich, I. (2019). Accounting for attention in sequential sampling models of decision making. *Current Opinion in Psychology*, *29*, 6–11. https://doi.org/10.1016/j.copsyc.2018.10.008

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, *3*, 193. https://doi.org/10.3389/fpsyg.2012.00193

Krucien, N., Ryan, M., & Hermens, F. (2017). Visual attention in multi-attributes choices: What can eye-tracking tell us? *Journal of Economic Behavior & Organization*, *135*, 251–267. https://doi.org/10.1016/j.jebo.2017.01.018

Krumpe, T., Gerjets, P., Rosenstiel, W., & Spüler, M. (2018). Decision confidence: EEG correlates of confidence in different phases of a decision task. *bioRxiv*, p. 479204.

Le, H. T., & Carrel, A. L. (2019). Happy today, satisfied tomorrow: emotion–satisfaction dynamics in a multi-week transit user smartphone survey. *Transportation*, *48*, 45–66. https://doi.org/10.1007/s11116-019-10042-6

Li, P., Markkula, G., Li, Y., & Merat, N. (2018). Is improved lane keeping during cognitive load caused by increased physical arousal or gaze concentration toward the road center? *Accident Analysis & Prevention*, *117*, 65–74. https://doi.org/10.1016/j.aap.2018.03.034

Li, W., Hu, X., Long, X., Tang, L., Chen, J., Wang, F., & Zhang, D. (2020). EEG responses to emotional videos can quantitatively predict big-five personality traits. *Neurocomputing*, *415*, 368–381. https://doi.org/10.1016/j.neucom.2020.07.123

Liu, Y., Sourina, O., & Nguyen, M. K. (2011). Real-time EEG-based emotion recognition and its applications. In *Transactions on computational science XII* (pp. 256–277). Springer.

Lizana, P., de Dios Ortúzar, J., Arellana, J., & Rizzi, L. I. (2021). Forecasting with a joint mode/time-of-day choice model based on combined rp and sc data. *Transportation Research Part A: Policy and Practice*, *150*, 302–316.

Lusk, J. L., Crespi, J. M., McFadden, B. R., Cherry, J. B. C., Martin, L., & Bruce, A. (2016). Neural antecedents of a random utility model. *Journal of Economic Behavior & Organization*, *132*, 93–103. https://doi.org/10.1016/j.jebo.2016.09.018

Mahieu, P.-A., Crastes, R., Louviere, J., & Zawojska, E. (2016). Rewarding truthful-telling in stated preference studies. Technical Report, Faculty of Economic Sciences, University of Warsaw.

Markkula, G., Uludağ, Z., Wilkie, R. M., & Billington, J. (2021). Accumulation of continuously time-varying sensory evidence constrains neural and behavioral responses in human collision threat detection. *PLOS Computational Biology*, *17*(7), e1009096. https://doi.org/10.1371/journal.pcbi.1009096

Merat, N., Jamson, A. H., Lai, F. C., & Carsten, O. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors*, *54*(5), 762–771. https://doi.org/10.1177/0018720812442087

Millen, A. E., & Hancock, P. J. (2019). Eye see through you! eye tracking unmasks concealed face recognition despite countermeasures. *Cognitive Research: Principles and Implications*, *4*(1), 1–14.

Mullett, T. L., & Stewart, N. (2016). Implications of visual attention phenomena for models of preferential choice. *Decision*, *3*(4), 231. https://doi.org/10.1037/dec0000049

Nowacki, J., Heekeren, H. R., Deuter, C. E., Joerißen, J. D., Schröder, A., Otte, C., & Wingenfeld, K. (2019). Decision making in response to physiological and combined physiological and psychosocial stress. *Behavioral Neuroscience*, *133*(1), 59. https://doi.org/10.1037/bne0000288

Nunes, P. A., & Schokkaert, E. (2003). Identifying the warm glow effect in contingent valuation. *Journal of Environmental Economics and Management*, *45*(2), 231–245. https://doi.org/10.1016/S0095-0696(02)00051-7

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 141–144).

Paschalidis, E., Choudhury, C. F., & Hess, S. (2018). Modelling the effects of stress on gap-acceptance decisions combining data from driving simulator and physiological sensors. *Transportation Research Part F: Traffic Psychology and Behaviour*, *59*, 418–435. https://doi.org/10.1016/j.trf.2018.09.019

Paschalidis, E., Choudhury, C. F., & Hess, S. (2019). Combining driving simulator and physiological sensor data in a latent variable model to incorporate the effect of stress in car-following behaviour. *Analytic Methods in Accident Research*, *22*, 100089. https://doi.org/10.1016/j.amar.2019.02.001

Pike, S., Kotsi, F., Oppewal, H., & Wang, D. (2020). Determining the importance of stopover destination attributes: Integrating stated importance, choice experiment, and eye-tracking measures. *Journal of Hospitality & Tourism Research*, *46*(1), 1096348020980818.

Płużyczka, M. (2018). The first hundred years: A history of eye tracking as a research method. *Applied Linguistics Papers*, *25*(4), 101–116. https://doi.org/10.32612/uw.25449354.2018.4.pp.101-116

Rao, L.-L., Zhou, Y., Xu, L., Liang, Z.-Y., Jiang, T., & Li, S. (2011). Are risky choices actually guided by a compensatory process? New insights from fMRI. *PLoS One*, *6*(3), e14756. https://doi.org/10.1371/journal.pone.0014756

Rasch, C., Louviere, J. J., & Teichert, T. (2015). Using facial emg and eye tracking to study integral affect in discrete choice experiments. *Journal of Choice Modelling*, *14*, 32–47. https://doi.org/10.1016/j.jocm.2015.04.001

Rodriguez, C. A., Turner, B. M., Van Zandt, T., & McClure, S. M. (2015). The neural basis of value accumulation in intertemporal choice. *European Journal of Neuroscience*, *42*(5), 2179–2189. https://doi.org/10.1111/ejn.12997

Ryan, M., Krucien, N., & Hermens, F. (2018). The eyes have it: Using eye tracking to inform information processing strategies in multi-attributes choices. *Health Economics*, *27*(4), 709–721. https://doi.org/10.1002/hec.v27.4

Sanders, N., Choo, S., & Nam, C. S. (2020). The EEG cookbook: A practical guide to neuroergonomics research. In *Neuroergonomics: principles and practice* (pp. 33–52). Springer.

Sanei, S., & Chambers, J. A. (2013). *EEG signal processing*. John Wiley & Sons.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322. https://doi.org/10.1038/nn1150

Smets, E., De Raedt, W., & Van Hoof, C. (2018). Into the wild: The challenges of physiological stress detection in laboratory and ambulatory settings. *IEEE Journal of Biomedical and Health Informatics*, *23*(2), 463–473. https://doi.org/10.1109/JBHI.6221020

Spinks, J., & Mortimer, D. (2015). 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), 1–13. https://doi.org/10.1186/s12911-016-0251-1

Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1228–1248. https://doi.org/10.1016/j.neubiorev.2012.02.003

Starcke, K., Ludwig, A.-C., & Brand, M. (2012). Anticipatory stress interferes with utilitarian moral judgment. *Judgment and Decision Making*, *7*(1), 61. https://doi.org/10.1017/S1930297500001832

Tarabay, R., & Abou-Zeid, M. (2021). A dynamic hybrid choice model to quantify stress in a simulated driving environment. *IEEE Transactions on Intelligent Transportation Systems*, *23*(7), 6390–6405. https://doi.org/10.1109/TITS.2021.3056146

Telpaz, A., Webb, R., & Levy, D. J. (2015). Using EEG to predict consumers' future choices. *Journal of Marketing Research*, *52*(4), 511–529. https://doi.org/10.1509/jmr.13.0564

Uggeldahl, K., Jacobsen, C., Lundhede, T. H., & Olsen, S. B. (2016). Choice certainty in discrete choice experiments: Will eye tracking provide useful measures? *Journal of Choice Modelling*, *20*, 35–48. https://doi.org/10.1016/j.jocm.2016.09.002

Vij, A., & Walker, J. L. (2016). How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological*, *90*, 192–217. https://doi.org/10.1016/j.trb.2016.04.021

Webb, R., Glimcher, P. W., Levy, I., Lazzaro, S. C., & Rutledge, R. B. (2013). Neural random utility and measured value. Available at SSRN, pp. 1–36.

Webb, R., Levy, I., Lazzaro, S. C., Rutledge, R. B., & Glimcher, P. W. (2019). Neural random utility: Relating cardinal neural observables to stochastic choice behavior. *Journal of Neuroscience, Psychology, and Economics*, *12*(1), 45. https://doi.org/10.1037/npe0000101

Wu, E. Q., Zhu, L. -M., Li, G. -J., Li, H. -J., Tang, Z., Hu, R., & Zhou, G. -R. (2021). Nonparametric hierarchical hidden semi-markov model for brain fatigue behavior detection of pilots during flight. *IEEE Transactions on Intelligent Transportation Systems*, *23*(6), 5245–5256. https://doi.org/10.1109/TITS.2021.3052801

Xu, J., Min, J., & Hu, J. (2018). Real-time eye tracking for the assessment of driver fatigue. *Healthcare Technology Letters*, *5*(2), 54–58. https://doi.org/10.1049/htl2.v5.2

Ziv, G. (2016). Gaze behavior and visual attention: A review of eye tracking studies in aviation. *The International Journal of Aviation Psychology*, *26*(3-4), 75–104. https://doi.org/10.1080/10508414.2017.1313096

Zysset, S., Wendt, C. S., Volz, K. G., Neumann, J., Huber, O., & von Cramon, D. Y. (2006). The neural implementation of multi-attribute decision making: A parametric fMRI study with human subjects. *Neuroimage*, *31*(3), 1380–1388. https://doi.org/10.1016/j.neuroimage.2006.01.017