**Article:**
Bogacz, M, Hess, S, Calastri, C et al. (2021) Modelling risk perception using a dynamic hybrid choice model and brain-imaging data: Application to virtual reality cycling. Transportation Research Part C: Emerging Technologies, 133. 103435. ISSN: 0968-090X

https://doi.org/10.1016/j.trc.2021.103435

# Modelling risk perception using a dynamic hybrid choice model and brain-imaging data: application to virtual reality cycling

Martyna Bogacz, Stephane Hess, Chiara Calastri, Charisma F. Choudhury, Faisal Mushtaq, Muhammad Awais, Mohsen Nazemi, Michael A.B. van Eggermond, Alexander Erath

## Abstract

Road risk analysis is one of the key research areas in transport, where the impact of perceived risk on choices, especially in a dynamic setting, has been long recognised. However, due to the lack of dynamic data and the difficulty in capturing risk perception, the existing studies typically resort to static and stated approaches to infer the experienced level of risk of individuals. In this paper, we aimed to address this research gap through developing a hybrid choice model that jointly employed dynamic data on cycling behaviour in virtual reality and neural data to evaluate how the fluctuations in momentary risk perception influence the behaviour of cyclists. The results of the developed model confirm our hypotheses, demonstrating that cyclists reduce their speed when approaching a junction as the potential for a collision with passing cars increases. Moreover, the latent component allowed us to establish a link between the neural data, the amplitude of alpha brainwaves, and objective risk measures. In line with our hypothesis, we found that decreased alpha amplitude is associated with higher perceived risk which in turn increases the likelihood of braking. The implications of our study are manifold. On the one hand, it shows the ability of virtual reality to elicit complex cyclists' behaviour and the feasibility of a joint collection of dynamic neural and choice data. On the other hand, we demonstrate the potential of the employment of neural data in a hybrid model framework as an indicator of risk that allows us to gain a better understanding of cycling behaviour and associated neural processing. These promising findings pave the way for future studies that would explore the advantages of neuroscientific inputs in the choice models.

**Keywords:** *cycling, virtual reality, EEG, hybrid choice model*

## 1 Introduction

Modelling risk perception in the transport context plays a critical role in the prediction of individuals' behaviour in a risky situation on the road. Traditionally, perceived risk and its effect on choices have been explored by means of paper or web-based stated preference (SP) surveys or simple laboratory experiments. In particular, direct questionnaires and attitudinal scales have been widely used in the transport context, for example, Ram and Chand (2016); Rundmo and Iversen (2004); Ulleberg and Rundmo (2003) to obtain insights on the perception of different modes, willingness to pay for their use (Román et al., 2014) as well as to assess road users' risk and safety perception (Zhang et al., 2011). Therefore, Driver Behaviour Questionnaire (Reason et al., 1990) or Young Driver Attitude Scale (Malfetti et al., 1989) are commonly used measures. They are claimed to perform well when respondents have sufficient experience or general knowledge about the context to reliably compare the situation in question, assess and report it (Mummolo and Peterson, 2019). Nonetheless, these explicit approaches have also been demonstrated to be susceptible to biases stemming from different

sources, such as experimenter effect (Iyengar, 2011), degree of familiarity with the topic (Patterson and Mattila, 2008), survey wording or social desirability bias (Näher and Krumpal, 2012).

Therefore, the major issue associated with these approaches is the possible incongruence between the stated and experienced risk and between the stated and actual actions. For instance, Andersson (2013) found a significant discrepancy between stated and observed willingness to pay for traffic safety, with Svensson (2009) also showing an inconsistency between the willingness to pay for risk reduction and the precautionary behaviour actually used to reduce it. Consequently, some studies attempted to resort to more implicit approaches to alleviate the issues associated with self-reported measures (Bennett and Vijaygopal, 2018). One of the examples is the Implicit Attitude Test (IAT) that measures the strength of association between concepts (e.g. man, woman) and evaluations (e.g. good, bad). It is based on the principle that the response time for the associations which are consistent with respondents' beliefs will be shorter than for the inconsistent ones. However, these implicit approaches are not problem less, where a study by Kim (2003) showed that the respondents can be successfully taught how to influence the IAT results. Beyond, De Houwer et al. (2007) demonstrated that experimenters were able to induce the formation of new attitudes in participants, which, in turn, had an impact on the outcome of the IAT measure. For this reason, instead of self-reported responses, which are prone to errors, neural or physiological activations can be measured to provide a raw, neural (or bodily) response to a specific situation.

Additionally, a limitation arises from the lack of dynamic data on risk, where attitudinal questions provide only a static indication of the inherently dynamic risk perception. While the experimental techniques may not capture the experience of risk to the same degree as real-life settings, some headway can be made with augmented and virtual reality (VR) technologies. These have been previously used, mostly in psychology research, to measure risk perception (Dixit et al., 2015) as well as physiological responses (Johnson et al., 2011; Shechtman et al., 2009; Chirico et al., 2017). A key advantage is the dynamic nature of the VR experiments, where many studies, focussed on risk, have adopted such a dynamic approach to improve realism and better capture the reactions to momentary changes in risk perception on the road (Frankenhuis et al., 2010; Mai, 2017; Underwood et al., 2011).

Virtual reality provides an alternative to the established methods of measuring risk perception and associated stress levels. Crucially for the present work, it also enables joint use with physiological devices and neuroimaging techniques such as electroencephalography (EEG), where the signal can be a proxy for risk perception. Previous literature in the transport context employed the EEG to predict driving behaviour on an operational level (Michon, 1985) including braking, acceleration and overtaking (Haghani et al., 2021). For instance, Kim et al. (2014) identified and used neural patterns of emergency braking situation to predict future braking. While other studies used biometric data to detect drivers' fatigue or drowsiness (Hu et al., 2009; Furman and Baharav, 2010). For example, a study by Awais et al. (2017) required the participants to drive in a simulator at a constant speed for a prolonged period to evoke drowsiness. Meanwhile, their electrical activity of the brain and heart rate data were recorded

to then make inferences about the relationship between these biometric measures and drivers' state. These studies by capturing the physiological and neural measures associated with different behaviours and states allow for a better understanding of mechanisms that drive them, which in turn allows to detect and predict them more accurately.

Therefore, a major opportunity arises with the greater accessibility to neuroimaging equipment to provide a new perspective on behaviour in risky and stressful conditions. Studies in neuropsychology have established that α waves[1] are a reliable stress marker (Verona et al., 2009; Lewis et al., 2007; Nishifuji et al., 2010; Vanitha and Krishnan, 2016; Seo and Lee, 2010) and links have been made between α wave variations and its role in attention and perception (Magosso et al., 2019; Ray and Cole, 1985; Cooper et al., 2006). In particular, a study by Brouwer et al. (2011) looked at neural correlates of stress evoked by the use of virtual reality. They found that simulated stressful conditions significantly influenced several physiological stress indicators used in the study, including frontal α power. Finally, Magosso et al. (2019) used virtual reality simulations and electroencephalography to detect changes in attention. The results of the study demonstrated that the α power decreased during tasks which required high attention to the external environment and conversely, α amplitude increased when the attentional demands of the task diminished. The results of these studies provide a promising basis for using α wave data as an indicator for the perceived risk, in the transport context, which increases the attentional demands. In particular, its usefulness in modelling of the perceived risk on the road could be explored if continuous EEG and behavioural data are collected jointly.

In a choice modelling context, and with perceived risk being unobserved, Hybrid Choice Models (HCM) are a suitable framework to incorporate neural data into the model. HCMs have been previously used mostly to combine latent constructs such as attitudes, opinions or perceptions together with observed choices in a single model structure by the means of measurable indicators (Ben-Akiva et al. (2002); Bolduc and Alvarez-Daziano (2010); Abou-Zeid and Ben-Akiva (2014)). More recently, Paschalidis et al. (2019) applied a hybrid framework with heart rate and skin conductance as indicators of unobserved stress to investigate its impact on driving behaviour.

The present paper addresses the shortcoming of past work by exploiting these novel opportunities in terms of an experiment that employs neuroimaging techniques to explore cycling risk perception in virtual reality simulation. We employ a dynamic HCM with an unobserved risk as a latent variable and we use α wave data as an indicator of the perceived risk to better explain the behavioural responses of the cyclist in the simulated environment. This approach allows us to achieve better understanding of cyclists' choices and make inferences about their neural background.

[1] Brain patterns form sinusoidal waves that are measured from peak to peak, with amplitude ranging from 0.5 to 100 μV. The brainwaves are measured in cycles per second (Hertz) which are also known as a frequency of brain wave activity. There are five major brainwaves identified: beta, alpha, theta, delta and gamma. The frequency of alpha brainwaves ranges between 8-13 Hz (Isa et al., 2014; Teplan et al., 2002; Ambekar and Achrekar, 2014).

The remainder of this paper is organised as follows. We present our specific hypotheses guided by the literature in the next section. The data collection design and sample characteristics are presented next, followed by the model structure and specification. We next turn to the results section, followed by the discussion section which reviews the insights from the analysis.

## 2 Hypotheses

In this section, we propose a number of hypotheses related to the changes in cyclists' behaviour with respect to fluctuations in the perceived level of risk due to dynamic traffic conditions and associated neural processing of their choices.

**Behavioural data**

- Hypothesis 1a: The presence of a passing car at the junction increases perceived risk, which in turn increases cyclists' propensity to reduce speed.

- Hypothesis 1b: The increase (or decrease) in perceived risk and hence the propensity to reduce (or increase) speed is a function of the remaining distance to reach the junction. That is, a shorter distance until the potential collision with the car at the junction increases cyclists' propensity to reduce speed.

These hypotheses are driven by the fact that junctions are parts of the road in our experiment where accidents with cars are possible, while speed reduction is the most common avoidance manoeuvre among cyclists (Johnson et al., 2010). Therefore, closer proximity to the junction and the presence of a car are expected to make cyclists brake more often in order to avoid collisions with other agents on the road or to gain more time to assess the dynamic situation at the junction, similarly to how we expect cyclists to react in reality when approaching a dangerous road.

**Neural data - α amplitude**

- Hypothesis 2: The α wave decreases (or increases) with increased (or reduced) risk.

This hypothesis draws on previous studies in the transport field which looked at the pre- diction and detection of braking intention based on the EEG features (Kim et al., 2014; Haufe et al., 2011; Hernández et al., 2018), providing evidence of a potential link between them. In particular, the existing neuroscientific literature links α wave to stress, visual attention, task performance and information processing relevant to the task (Verona et al., 2009; Fournier et al., 1999; Borghini et al., 2014; Fairclough et al., 2005; Slobounov et al., 2000). In particular, multiple previous studies have shown that the changes of α activity were related to the strength of attention to external stimuli required by the task (Mann et al., 1996; Vanni et al., 1997; Lei and Roetting, 2011; Simon et al., 2019). In other words, the presence of a dangerous element in the scenario (car crossing at the junction) is considered to increase the task complexity, and consequently, higher attentional demands lead to a decrease in the α power (amplitude). Therefore, it is hypothesised that α power is negatively associated with perceived risk.

## 3 Experimental design and sample information

This section explains the experimental procedure, technical details of the data collection and describes the equipment employed in the experiment i.e. the instrumented bicycle, virtual reality goggles and the EEG device. Beyond, we present the key sample characteristics.

Following the experimental design from Bogacz et al. (2020) the study employed six scenarios with an immersive simulation of a street from the cyclist perspective. All the scenarios were similar with respect to certain aspects. Firstly, the cyclist was riding on the pavement, separated from the pedestrian zone on the left by a painted lane and segregated from the road traffic on the right by a grass strip, as seen in Figure 1.

*Figure 1: The location of the bicycle lane in the scenario.*


Secondly, in each scenario, there were two junctions that were places of a potential collision with cars, as demonstrated in Figure 2. Moreover, all the scenarios involved both, pedestrians and cars. The pedestrians could only walk along the bicycle lane, but cars could turn at the junctions. While the proportion of cars that would turn was constant across all the scenarios, the movement of a particular vehicle was random. Consequently, there were some differences between scenarios in terms of the actual number of cars that were present at the junction at the same time as the cyclist. This was also influenced by the cycling speed of the participant which, in turn, had an impact on the time when the cyclist arrived at the junction. The microsimulation of traffic used in the scenarios was designed using PTV Vissim software (PTV-Group, 2021) and later developed in Unity as a 3D, virtual reality scenario (Unity, 2017). The scenarios included a 360-degree view of the street around the cyclist and were responsive to their head movements. The sound was also included in the simulation to account for the auditory cues that are normally available to cyclists in real life. Importantly, the loudness of the vehicles depended on their distance to the cyclist so that the noise of a car increased as it was becoming closer (Bogacz et al., 2020: 247).

*Figure 2: An example of collision point - a junction.*


The reason for using six scenarios is the simultaneous collection of neural data that required more trials to achieve higher EEG signal stability. The experimental task was to cycle through the scenario at the desired speed until arriving at the finish line. The participants used the pedals to move forward and the hand brake located on the right-hand side of the handlebar to brake but they could not steer. The instrumented bike can be observed in Figure 3 and it is a part of the Future Cities Laboratories in Singapore.

*Figure 3: Instrumented bicycle used in the experiment*

The experimental session began with familiarising the participant with the operation of an instrumented bicycle, where all the participants could ride a bike in reality. Next, the participant was sat on the bicycle and the HTC Vive head-mounted display (HMD) (Borrego et al., 2018) and the EEG recording device Enobio (Riera et al., 2008) were put on their head. The Enobio headset used the following electrodes: FP1, FP2, Fz, C3, Cz, C4, P3 and P4 to record the signal

from the whole scalp at sampling rate of 128 Hz[2]. The compact build of the mobile EEG device enabled us to use it jointly with HTC Vive. To collect the baseline activity of the brain, the participants were asked to focus their gaze on one point on the screen for one minute. Subsequently, the same procedure was performed with eyes closed. It was followed by the trial session to familiarise the participants with the operation of the bicycle as well as the simulation. The summary of the data collected in the experiment is presented in Table 1.

*Table 1: Summary of data collected in the experiment.*

| Data type | Variable | Resolution of recording | Unit of measurement |
|---|---|---|---|
| Behavioural | Speed | 4 Hz | km/h m/s2 |
| Neural | Acceleration | 4 Hz | μV (microvolt) |
| Behavioural | α wave amplitude | 128 Hz | km/h m/s2 |

The sample consisted of fifty individuals recruited from staff and students of the National University of Singapore (NUS) as well as members of the general public. Nonetheless, two of them did not complete the whole experiment hence they were dropped from the analysis leading to a final sample size of 48. The mean age of the sample was 26.5 years (6.7 years standard deviation). Importantly, small sample size is a typical issue faced by the studies that use virtual reality or driving simulators (Camara et al., 2021; Nuñez Velasco et al., 2021) due to longer experiment duration and higher costs as compared to stated preference studies.

## 4 Methods for EEG data cleaning and extraction

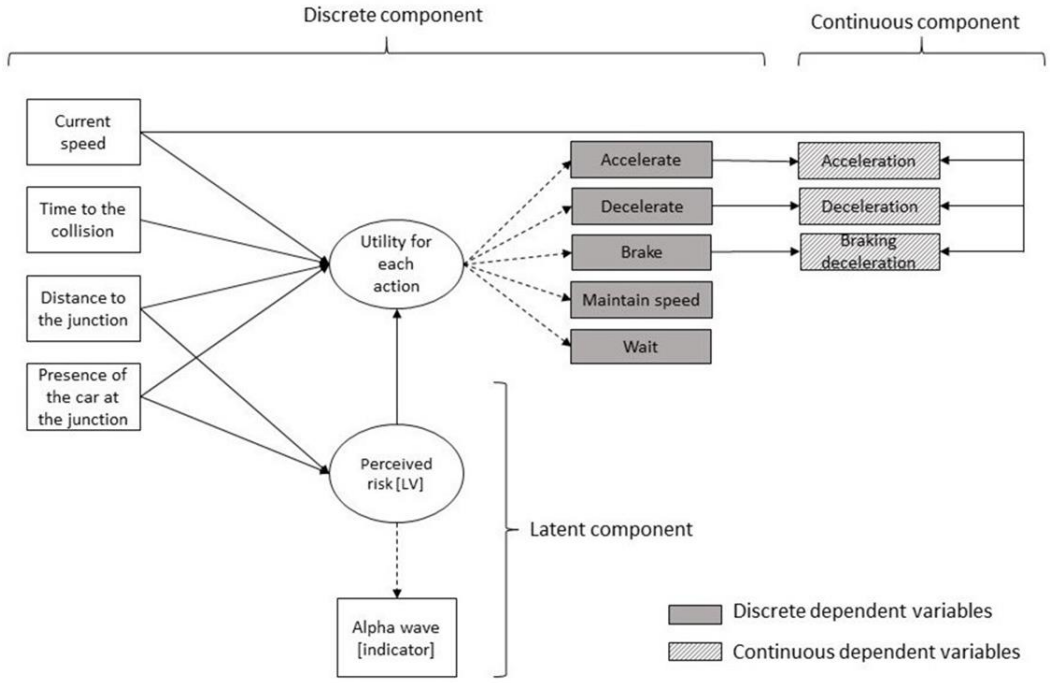
We used the EEG data cleaning protocol described in Bogacz et al. (2020), where firstly the 1-20 Hz bandpass filtering (BPF) was applied to preserve the components of the data within this specific band of frequencies (Christiano and Fitzgerald, 2003) and remove those which we were not interested in. The use of BPF allowed us to eliminate, for example, the noise stemming from external sources such as the operation of electronic equipment or internal, physiological sources. Then, the artefacts such as eye blinks and other body movements were removed manually using multiple source analysis method within BESA 6.0 software (MEGIS Software GmbH, Gräfelfing, Germany) (Berg and Scherg, 1994; Ille et al., 2002). Next, Welch's method was used to compute the power spectrum of the EEG (Welch, 1967), before we averaged the EEG data from 128 Hz to 4 Hz to match the resolution of the behavioural data. Finally, we normalized the logarithm of the α wave so that for each individual α values have a mean of 0 and a standard deviation of 1.

[2] Hertz (Hz) is a unit of temporal frequency which denotes the number of occurrences of an event per one second. For example, a recording resolution of 128 Hz means that the data has been collected 128 times per one second.

# 5 Modelling framework

In this section, we describe the modelling framework used in this study, where we developed a hybrid choice model with discrete and continuous choice components and a latent variable. Cyclists' choices of action in the virtual environment were recorded every quarter of a second. In Figure 4, the latent variables are presented in ovals and the observed variables are presented in rectangles. The utility of each of the discrete actions (i.e. accelerate, brake, decelerate, maintain speed and wait) is influenced by the different exogenous variables related to the situation in the scenario and latent perceived risk. The latent component of the model consists of the latent risk with α wave as its indicator and two explanatory variables. If the cyclist chooses to accelerate, brake, or decelerate, then the continuous component for a given action is considered. Importantly, in the continuous part, the acceleration is split into two cases, namely acceleration when previously stopped and when previously in motion. We present a more detailed description of each model component in turn.

*Figure 4: A model structure*




## Specification of latent risk

The developed hybrid choice model encompasses a latent component that seeks to capture the impact of risk on cyclist's actions, where it is believed that in real life the individuals also adjust their cycling in response to the observed level of risk. We assumed that the perceived risk is influenced by a cyclist's distance to the junction and the presence of a car at the junction. The latter was included as an additive interaction term. The structural equation for the latent variable can be seen in Equation 1.

$$\theta_{latent-risk} = (\gamma_{dist-to-junction} + \gamma_{car-present} \cdot x_{car-present}) \cdot x_{dist-to-junction} + \varepsilon_{\alpha} \quad (1)$$

Where:

- $x_{dist-to-junction}$ is the variable representing the cyclist's distance to the junction (measured in meters);
- $x_{car-present}$ is a dummy variable, taking value 1 if a car is present at the junction at the current point in time and 0, otherwise;
- $\varepsilon_{\alpha}$ is an error term.

Further, the corresponding parameters:

- $\gamma_{dist-to-junction}$ is the parameter representing the influence of the cyclist's distance to the junction on the perceived risk;
- $\gamma_{car-present}$ is the interaction parameter representing the additive shift in the perceived risk if a car is present at the junction.

Importantly, different forms of structural equations for the latent variable were tested, including, for example, discrete levels of time to collision, however, the final specification was chosen as described above because the analysis of the data showed that the level of risk that the cyclist perceives in each time point is related to their distance to the hazardous part of the scenario, i.e. the junction. This particular relationship can possibly be explained by the fact that distance is an easier feature to be judged by the cyclist compared to, for example, time until arrival to the junction. Then, we included the presence of the car at the junction which undoubtedly amplifies the perceived risk since it materialises the potential hazard. Moreover, to obtain more insights into the relationship between variables included in the structural equation and the indicators, one-sided t-test was conducted, firstly on mean values of alpha amplitude in cases when the car was absent and present at the junction. The result was in line with our expectations, where α amplitude was significantly lower (t-ratio = 1.92) in cases where there was a car present at the junction. Moreover, Spearman's correlation coefficient of 0.16 between α amplitude and distance to the junction is statistically significant and although it is small, it suggests a positive relationship between these two variables, as expected.

**Specification of discrete component**

***Dependent variable***

The choice variable for the discrete part of the model allows for five possible actions, namely, *accelerate, brake, decelerate, maintain speed,* and *wait*. These actions are linked to the changes in cycling speed, where *accelerate* refers to an increase in speed, *brake* encompasses cases of an abrupt decline in speed, while *decelerate* covers instances where a person slows down at a slower rate than when braking. This distinction was made to account for different underlying reasons for these actions. For example, braking may be the result of a sudden emergence of a

car at the junction leading to an increase in the risk of collision; while deceleration may be chosen when there are no imminent risks of collision, but the need for increased caution (e.g. when the cyclist is still far from the junction) or it is possible that deceleration arises simply when the cyclist stops or reduces pedalling due to physical exhaustion. Cases, where a person is moving while keeping a constant speed, are classified as *maintain speed*. Finally, *wait* refers to instances where a cyclist has stopped and remains still. Importantly, the *wait* is not available when the cyclist is currently moving. On the other hand, when the cyclist is not in motion, the only two available actions are *wait* and *accelerate*. The *wait* was included as an alternative because we observed in the data that cyclists chose it close the junction to allow for the traffic to clear and cross the junction at an appropriate time.

To summarise, when the cyclist is moving, the available actions are: *accelerate, brake, decelerate, maintain speed*, whereas when the cyclist is not moving, he can either *wait* or *accelerate*. For this reason, the reference category in the model is *accelerate* because this is the only action that is always available, whether moving or not. Finally, this categorisation of actions was chosen because it allows us to capture a cyclist's reactions to the dynamic environment in the scenario by the degree to which he adjusts his current speed, similarly to individuals cycling in the real world. In other words, we are interested in the instances when the changes in the external environment induce the cyclist to switch his currently chosen action to a different one.

***Utility functions***

As a result of our approach described in the previous section, the utility associated with each action is a function of:

- the cyclist's distance to the junction;
- the cyclist's speed;
- the presence of a car at the junction and the time until the collision with the given car (if there is one), considering the current speed of the cyclist.

In order to account for the potential non-linear effects of distance and speed, we included the polynomials of distance to the junction and cyclist's speed up to the third order. Furthermore, the impact of a car being present at the junction is captured by a dummy which indicates a potential collision with a car, given the cyclist's current distance to the junction and his current speed, which are recorded every quarter second. The reason for using the attribute on a possible collision in the future is that it allows us to capture how a cyclist adjusts his current behaviour to avoid colliding with the car. We did not use the presence of the car at the junction, because, while it increases the perceived risk, it does not automatically require acting upon. In other words, it would not enable us to observe how a cyclist changes his behaviour as a result of an increase in situation riskiness. Moreover, we included a variable that accounts for the time remaining until collision (again with the resolution of a quarter second) if a car is present at the junction. This time component allows us to consider the impact of the remaining time to the

collision on the choice of different actions. Finally, we also included a random variable to account for individuals' heterogeneity in preferences for different actions.

Therefore, below we present the utility functions associated with the decision of a cyclist *n* to choose one of the five actions (Acc = accelerate, Br = brake, Dec = decelerate, Maintain = maintain speed, Wait = wait) at time *t*, where accelerate is used as the baseline. Importantly, Equations 2 - 5 show full specification but some components were removed during the actual estimation of the model because they were insignificant, or they were not relevant explanators for that given alternative[3]. For instance, in the actual utility function for Wait, we did not retain the speed or distance to the junction as they were not meaningful components when the cyclist was not moving.

$$
\begin{aligned}
V_{Maintain_{n,t}} = {} & \delta_{Maintain} + \sigma_{Maintain} \cdot \xi_{Maintain} \\
& + \beta_{dist-to-junction_{Maintain}} \cdot x_{dist-to-junction_{n,t}} \\
& + \beta_{dist-second_{Maintain}} \cdot x^2{}_{dist-to-junction_{n,t}} \\
& + \beta_{dist-third_{Maintain}} \cdot x^3{}_{dist-to-junction_{n,t}} \\
& + \beta_{speed_{Maintain}} \cdot x_{speed_{n,t}} + \beta_{speed-second_{Maintain}} \\
& \cdot x^2{}_{speed_{n,t}} + \beta_{speed-third_{Maintain}} \cdot x^3{}_{speed_{n,t}} \\
& + \beta_{collision_{Maintain}} \cdot x_{collision_n} \cdot x_{time-collision_n}^{\lambda_{time-collision_{Maintain}}} \\
& + \beta_{risk_{Maintain}} \cdot \theta_{latent-risk}
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
V_{Br_{n,t}} = {} & \delta_{Br} + \sigma_{Br} \cdot \xi_{Br} + \beta_{dist-to-junction_{Br}} \cdot x_{dist-to-junction_{n,t}} \\
& + \beta_{dist-second_{Br}} \cdot x^2{}_{dist-to-junction_{n,t}} + \beta_{dist-third_{Br}} \\
& \cdot x^3{}_{dist-to-junction_{n,t}} + \beta_{speed_{Br}} \cdot x_{speed_{n,t}} \\
& + \beta_{speed-second_{Br}} \cdot x^2{}_{speed_{n,t}} + \beta_{speed-third_{Br}} \\
& \cdot x^3{}_{speed_{n,t}} + \beta_{collision_{Br}} \cdot x_{collision_n} \cdot x_{time-collision_n}^{\lambda_{time-collision_{Br}}} \\
& + \beta_{risk_{Br}} \cdot \theta_{latent-risk}
\end{aligned}
\tag{3}
$$

[3] Beyond, there were other variables such as cyclist's distance to the junction after crossing it, which were tested as well but not included in the final models due to non-intuitive sign and statistical insignificance.

$$V_{Dec_{n,t}} = \delta_{Dec} + \sigma_{Dec} \cdot \xi_{Dec} + \beta_{dist-to-junction_{Dec}} \cdot x_{dist-to-junction_{n,t}} + \beta_{dist-second_{Dec}} \cdot x^2{}_{dist-to-junction_{n,t}} + \beta_{dist-third_{Dec}} \cdot x^3{}_{dist-to-junction_{n,t}} + \beta_{speed_{Dec}} \cdot x_{speed_{n,t}} + \beta_{speed-second_{Dec}} \cdot x^2{}_{speed_{n,t}} + \beta_{speed-third_{Dec}} \cdot x^3{}_{speed_{n,t}} + \beta_{collision_{Dec}} \cdot x_{collision_n} \cdot x_{time-collision_n}^{\lambda_{time-collision_{Dec}}} + \beta_{risk_{Dec}} \cdot \theta_{latent-risk} \quad (4)$$

$$V_{Wait_{n,t}} = \delta_{Wait} + \sigma_{Wait} \cdot \xi_{Wait} + \beta_{dist-to-junction_{Wait}} \cdot x_{dist-to-junction_{n,t}} + \beta_{dist-second_{Wait}} \cdot x^2{}_{dist-to-junction_{n,t}} + \beta_{dist-third_{Wait}} \cdot x^3{}_{dist-to-junction_{n,t}} + \beta_{speed_{Wait}} \cdot x_{speed_{n,t}} + \beta_{speed-second_{Wait}} \cdot x^2{}_{speed_{n,t}} + \beta_{speed-third_{Wait}} \cdot x^3{}_{speed_{n,t}} + \beta_{collision_{Wait}} \cdot x_{collision_n} \cdot x_{time-collision_n}^{\lambda_{time-collision_{Wait}}} + \beta_{risk_{Wait}} \cdot \theta_{latent-risk} \quad (5)$$

$$V_{Acc_{n,t}} = 0 \quad (6)$$

The $\delta_i$ parameters in Equations 2 to 5 above represent the alternative specific constants (ASC), where the subscripts refer to each action. We allowed for random heterogeneity in these preferences through the additional terms $\sigma_i$, which multiply a standard Normal variate $\xi_i$.

Also, there are several other components that look at the impact of different variables in the scenarios on the utilities, as seen in Table 2.

The Equations 2 to 5 also have parameters for each action, where for the ease of notation we use subscript *i* in the text that can denote *Accelerate, Brake, Decelerate, Maintain speed* and *Wait*:

- $\delta_i$ represent the alternative specific constant (ASC) for each action;
- $\sigma_i$ captures the heterogeneity in the alternative specific constant (ASC) for each action;

• $\beta_{dist-to-junction_i}$ , $\beta_{dist-second_i}$ and $\beta_{dist-third_i}$ are parameters which represent the impact of the distance to the junction on the utility for each action, first, second and third order, respectively;

• $\beta_{speed_i}$ , $\beta_{speed-second_i}$ and $\beta_{speed-third_i}$ are parameters which represent the impact of the cyclist's speed on the utility for each action, first, second and third order, respectively;

• $\beta_{collision_i}$ are the parameters which show the impact of the potential presence of the car at the junction on the utility for each action;

• $\beta_{risk_i}$ are the parameters that represent the impact of latent risk on different actions;

• $\lambda_{time-collision_i}$ are the parameters which represent the impact of the time left until the collision on each action.

*Table 2: Variables used in the utility functions.*

| Variable name | Variable meaning | Units |
|---|---|---|
| $x_{dist-to-junction}$ | cyclist's distance to the junction; used also in the polynomial form: $x_{dist-to-junction}2$ and $x_{dist-to-junction}3$ to capture the non-linearimpacts | metres |
| $x_{speed}$ | cyclist's speed; used also in the polynomial form: $x_{speed}2$, $x_{speed}3$ to capture the non-linear impacts | Km/h |
| $x_{collision}$ | dummy variables which denote if there would be a collision with the car at the junction when the cyclist would arrive there, given his current speed and distance to the junction | equal to 1 if true, 0 otherwise |
| $x_{time-collision}$ | remaining time to the collision between a cyclist and the car given cyclists current speed and distance to the junction | seconds |
| $\theta_{latent-risk}$ | latent risk variable | - |
| $\xi_i$ | a random variable that represents the individual's heterogeneity for different actions | - |

**Specification of the continuous component**

The continuous component of the model was developed to give an additional level of detail on cycling behaviour, providing information on the magnitude of the actions chosen in the discrete component (Varotto et al., 2018; Koutsopoulos and Farah, 2012). This not only allows to achieve a higher level of understanding of modelled behaviour, but it also helps to validate further the adopted methodological approach. Importantly, the aim of the continuous component was not to improve the model efficiency, but to gain behavioural insights, and from the mathematical standpoint it does not influence the performance of the model as the discrete and continuous components were estimated separately (see Appendix B, Table 5 for the full

output of the model without the continuous component). Therefore, in the continuous component of the model, the dependent variables are the continuous values of accelerate, decelerate and brake (called acceleration, deceleration and braking deceleration, respectively, in the continuous part), chosen by the cyclist. The model assumes that the cyclist firstly chooses one of these three actions in the discrete part of the model and subsequently he decides on the acceleration/deceleration associated with the chosen decision in the continuous part. The continuous decision is thus conditional on the discrete choice. In the case of acceleration in the continuous part, it differentiates between accelerating when the cyclist was moving in the previous time point and when he was waiting (i.e. was stopped). For the former, the current cycling speed is expected to affect the chosen value of acceleration while for the latter, the current speed is zero. Importantly, the dependent variables in all four cases are derived from the absolute values of acceleration but for each of them, the values are only considered when the given action is chosen in the discrete part, as described above. Therefore, for correct interpretation of the parameter estimates, it should be noted that the acceleration refers to increasing velocity in the case of accelerate as a discrete choice and decrease in velocity for deceleration and braking deceleration. Guided by the shapes of the distributions of the observed acceleration and deceleration values, the dependent variable is assumed to follow a log-normal distribution with the mean being a function of explanatory variables (e.g. current speed). Consequently, the specification of the continuous variable can be seen in Equation 7.

$$P\left(log\left(Acc_{n,t}\right)\right) = \frac{1}{\left(\sigma_{log\left(Acc_{n,t}\right)}\sqrt{2\pi}\right)} * exp\left(-\frac{1}{2}\left(\frac{\left(log(Acc) - x_{Speed_{n,t}} - \mu_{log\left(Acc_{n,t}\right)}\right)}{\sigma_{log\left(Acc_{n,t}\right)}}\right)^2\right) \quad (7)$$

Where,

- $\mu_{log\left(Acc_{n,t}\right)}$ denotes the mean of the logarithm of the dependent variables
- $\sigma_{log\left(Acc_{n,t}\right)}$ denotes the standard deviations of the logarithm of the dependent variables
- $x_{Speed_{n,t}}$ is the cyclist's current speed.

**Specification of the measurement component**

The final component of the hybrid structure is the measurement model which is used to link the latent variable with its indicators. In our model, the risk is not observed or measured directly, but instead, it is manifested in its indicator, the alpha (α) brain activity, derived from the EEG data. To account for the heterogeneity in the α ranges across the respondents, normalized values of α have been employed (as used by Paschalidis et al. (2019) for physiological indicators and Makeig and Jung (1995) for the EEG signal). Multiple previous studies showed that the changes of α activity were related to the strength of attention to external stimuli required by the task, where naturally more hazardous elements increase these attention demands. The measurement model can be seen in Equation 8.

$$\alpha = \zeta_{\alpha}\theta_{latent-risk} + \nu_{\alpha} \quad (8)$$

Where,

• $\theta_{latent-risk}$ is latent risk variable;

• $v_{\alpha}$ is an independent error term, and the corresponding parameter:

• $\zeta_{\alpha}$ is a parameter relating the latent risk to its indicator, the α wave.

# 6 Results

This section presents the full results of the hybrid model in Table 3. Different components of the modelling approach developed in this paper are described in turns. The model was estimated simultaneously with 100 Halton draws using the Apollo software (Hess and Palma, 2019).

*Table 3: A discrete-continuous model (robust standard errors and t-ratios in brackets).*

LL (final, whole model): -149642.40
AIC: 299370.70
BIC: 299752.40
LL (final, MNL): -45129.09
LL (final, acceleration when static): 464.95
LL (final, acceleration when moving): -14384.50
LL (final, deceleration): -7492.83
LL (final, braking deceleration): -8338.87
LL (final α): -74790.59
Number of observations = 52896
Number of respondents = 48

| | | Action | Estimate (rob.t-ratios) |
|---|---|---|---|
| ASC ($\delta$) | | Maintain | -0.2644 (0.0797; -3.32) |
| | | Brake | -2.4598 (0.1256; -19.58) |
| | | Decelerate | -2.9847 (0.1183; -25.23) |
| | | Wait | 4.1069 (0.1869; 21.98) |
| | | Accelerate | 0 |
| Individual's heterogeneity ($\beta_{\sigma}$) | | Maintain | 0.2241 (0.0217; 10.34) |
| | | Brake | 0.0363 (0.0397; 0.92) |
| | | Decelerate | -0.1051 (0.0445; -2.36) |
| Collision | | Maintain | 0.2031 (0.0384; 5.28) |
| | | Brake | 0.7535 (0.1542; 4.89) |
| | | Decelerate | 0.3616 (0.1185; 3.05) |
| | | Wait | 0.6292 (0.1758; 3.58) |
| Time to collision | | Brake | -0.5280 (0.0905; -5.83) |
| | | Decelerate | -0.1612 (0.1508; -1.07) |
| Distance to junction | 1st order | Maintain | 0.0444 (0.0038; 11.68) |
| | 2nd order | Maintain | -0.0835 (0.0085; -9.86) |
| | 1st order | Brake | 0.1900 (0.0138; 13.78) |
| | 2nd order | Brake | -0.2708 (0.0604; -4.48) |
| | 3rd order | Brake | 0.0195 (0.0080; 2.44) |
| | 1st order | Decelerate | 0.0850 (0.0125; 6.78) |
| | 2nd order | Decelerate | -0.2446 (0.0621; -3.94) |
| | 3rd order | Decelerate | 0.0166 (0.0078; 2.12) |

| | | | |
|---|---|---|---|
| Speed | 1st order | Maintain | 0.0605 (0.0062; 9.80) |
| | 2nd order | Maintain | -0.0026 (0.0004; -5.79) |
| | 1st order | Brake | 0.0458 (0.0141; 3.26) |
| | 2nd order | Brake | 0.0014 (0.0006; 2.25) |
| | 1st order | Decelerate | 0.1931 (0.0270; 7.16) |
| | 2nd order | Decelerate | -0.0058 (0.0019; -3.09) |
| | 3rd order | Decelerate | 0.0001 (0.0000; 3.06) |
| Continuous acceleration when stopped | | $\mu$ | -3.0631 (0.1386; -22.10) |
| | | $\sigma$ | 2.2931 (0.0795; 28.84) |
| Continuous acceleration when in motion | | $\mu$ | -0.3722 (0.0336; -11.09) |
| | | $\sigma$ | 1.0912 (0.0144; 75.58) |
| | | Speed | 0.0518 (0.0022; 23.90) |
| Continuous deceleration | | $\mu$ | -0.2864 (0.0435; -6.59) |
| | | $\sigma$ | 1.1047 (0.0196; 56.38) |
| | | Speed | 0.0394 (0.0026; 15.15) |
| Continuous braking deceleration | | $\mu$ | -0.4394 (0.0639; -6.87) |
| | | $\sigma$ | 1.1958 (0.0253; 47.32) |
| | | Speed | 0.0602 (0.0038; 15.97) |
| Latent risk (θ) | | Brake ($\beta_{risk_{Br}}$) | 0.3285 (0.0774; 4.24) |
| | | Distance to the junction ($\gamma_{dist-to-junction}$) | -0.0762 (0.0058; -13.11) |
| | | Car at the junction ($\gamma_{car-present}$) | 0.0041 (0.0015; 2.72) |
| | | ζα | -0.0937 (0.0187; -5.01) |

## Latent component

We start by looking at the latent variable since it is a key component of the model, where we see that higher risk significantly increases the propensity to brake (estimate = 0.3285), whereas for the other actions the effect is not significantly different from that to accelerate. The result for brake is plausible because higher risk induces more braking.

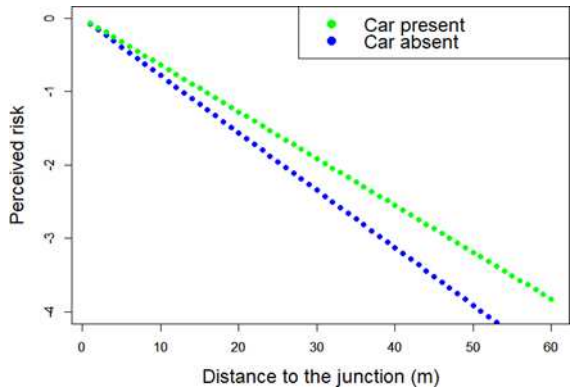
## Measurement model

Then, looking at the results of the measurement model, we observe a negative and significant $\zeta_\alpha$ coefficient, which suggests that increased risk decreases the amplitude of the α wave, which is in line with the existing literature.

## Structural model

Furthermore, in the structural model within the latent component, we observe a negative relationship between distance to the junction and perceived risk (estimate = -0.0762), which again makes sense where a higher distance to the hazardous area decreases perceived risk. Then, the presence of the car at the junction dummy variable is an additive shift for when there is a car present at the junction regardless of the cyclist's distance to the junction. The effect (of a car being present) on risk perception is visualised in Figure 5, where we see that the presence of the car results in an upward shift in the perceived risk. To clarify, the negative sign of the perceived risk is not relevant here because it only shows us the relationship within the latent construct where the plot allows us to understand how the explanatory variables (here the distance and presence of the car) influence the unobserved risk.

*Figure 5: Risk perception plot.*




**Discrete component**

We then turn to the discrete component of the model, where the ASCs show us the impact of different actions on utility, all else being equal, nevertheless, their significance suggests there are some elements omitted in the model that have an impact on the utility. Next, the σ estimated for maintain speed and decelerate are statistically significant showing the individual's heterogeneity for these two actions, while σ for wait and brake was not significantly different from zero. It suggests that individuals differ in their propensity to maintain speed and decelerate regardless of their perceived risk, which may be related to other factors such as the level of physical fitness, personal traits or cycling style.

Next, the dummy collision variables demonstrate that the potential presence of a car at the junction when cyclists arrive at the junction increases the propensity to brake. The next more likely actions are wait and decelerate and finally maintain speed (all relative to accelerate, which is, therefore, the least likely action). These again show realistic cycling behaviour which we would observe in real-life situations when cyclists approach a junction in the presence of a car.

Moreover, the time to collision parameters for brake and decelerate allowed us to factor in the time left to the collision given a cyclist's speed and distance to the junction. These show us that less time left to the collision increases the probability of braking and decelerating relative to accelerating. Again, these results are plausible where closeness to a potential accident with a car at the junction increases the likelihood of reaction by decreasing the speed. Further, for maintain speed and wait, the effect was not significantly different compared to accelerate.

Then, we see a number of distances to junction variables, where the results of these variables are demonstrated in Figure 6. We can see the graphs of probability for accelerate, decelerate, maintain speed and brake depending on the cycling speed and distance to the junction. The same graph is shown from different angles to make the interpretation easier.

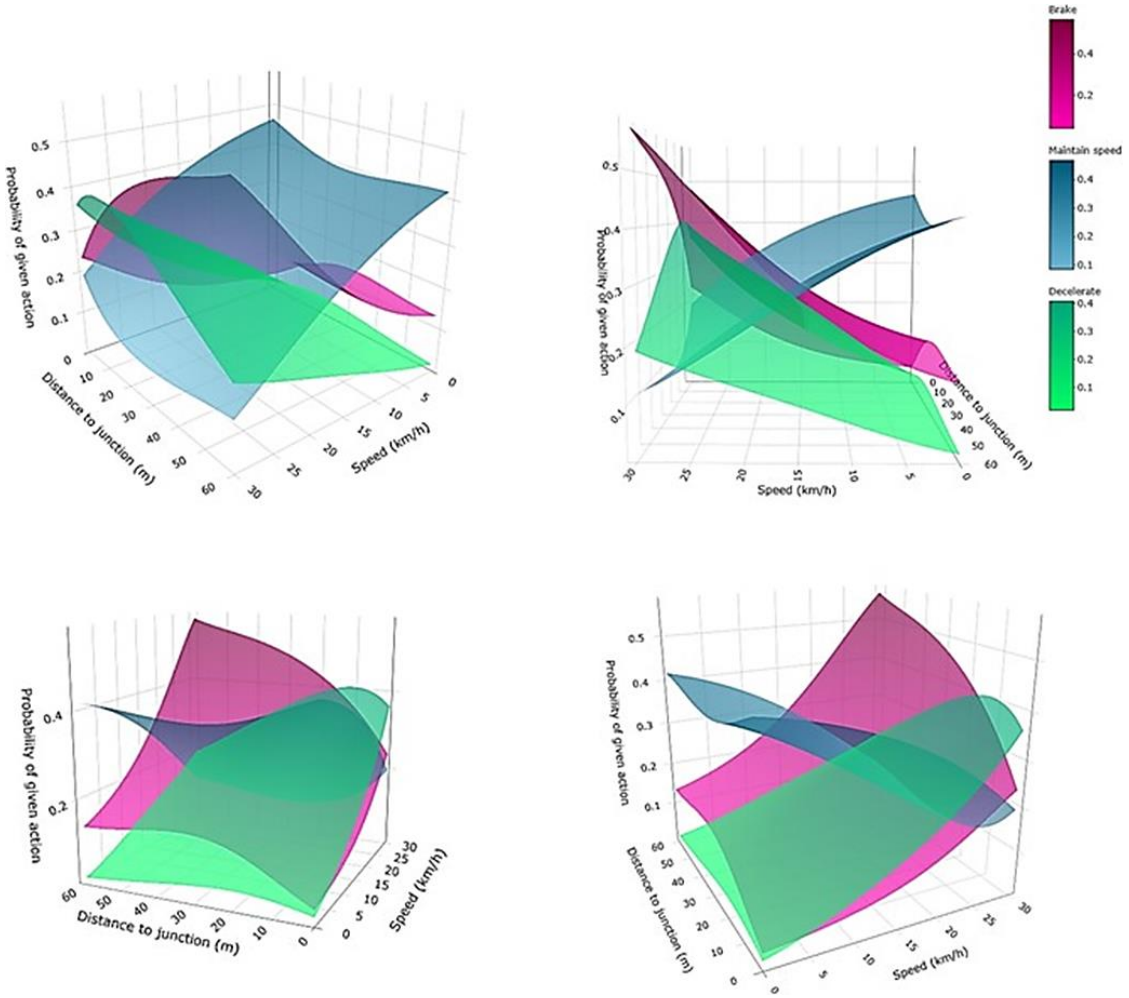
Looking at the graph for brake, we observe that the probability of this action is the highest at high-speed levels and decreases as the speed falls. It suggests that at a fast pace the cyclists are

ready to react quickly (by braking) to the changes in the environment. Moreover, at low speed, the probability of braking increases up to about 20 meters away from the junction which is an area where cyclists tend to slow down (or stop) to assess the situation at the junction. After this point, the probability of braking decreases sharply as the cyclists intend to cross the junction.

Next, looking at the graph for decelerate, we see that the probability of decelerating is the highest at the high-speed levels where there is the biggest scope for physical exhaustion to emerge, and it falls with speed. Similar to brake, the probability for decelerate increases up to about 15 meters from the junction.

Finally, looking at the graph for maintain speed, we see that the utility for this action is the highest when cyclists are going at low to medium speed levels and decreases as the cycling speed goes up, nearly independently of the distance to the junction. Although, at a high-speed level an increase in the probability for this action can be observed, as the cyclists become very close to the junction, suggesting cyclists' confidence to cross after assessing the situation at the junction.

*Figure 6: 3D graphs of the probability for each action given cyclist's current speed and distance to the junction.*



**Continuous component**

Finally, the continuous part of the model shows us that all the estimates are significant. Firstly, comparing the values of $\mu$ for acceleration when static and moving we see a considerably more negative value in case of stationary acceleration. This suggests that, on average, the acceleration rates are smaller when cyclists only start accelerating after being stopped, than when already moving, which can be expected as cyclist is only gaining speed from a standstill.

Next, looking at the values of $\sigma$ for acceleration when static and when in motion, we see that the standard deviation of acceleration from static position is twice as high, which can be attributed to the higher variability of individual strength input when starting the bicycle. Whereas when moving, the level of force required to accelerate is more homogenous because individuals benefit from momentum leading to smaller standard deviation of acceleration. Notably, the $\sigma$ for acceleration when previously moving, deceleration and braking deceleration are similar in magnitude.

Further, the estimated parameters for speed are positive in all three cases (it has been dropped from the model for acceleration when previously waiting because in this case the cycling speed is zero) and they show that at higher current speed, the magnitude of acceleration changes more. In the case of braking deceleration and deceleration, it shows that the scope for reducing the speed is larger if cyclist is going at a higher speed. While, in case of acceleration, we see that it becomes easier to accelerate.

**Hybrid choice model efficiency gain**

Finally, to assess the impact of the inclusion of the neural data in the hybrid model, we compare the model results between our model as showed in Table 3, and a reduced model, where the measurement component in the latent part was dropped (see Appendix A, Table 4 for the full model output). Hence, the role of the latent variable is explained only by the choice data. We, then, see that removing a measurement model changes the magnitude of distance variable in the structural equation and it becomes less significant (change from t-ratio = -13.11 to t-ratio = -1.72). Moreover, in the reduced model, the standard errors of variables increased from 0.0058 to 0.0305 for distance variable and from 0.0015 to 0.0021 for car presence variable. Therefore, the inclusion of the measurement model not only increases the efficiency of the parameters, but also confirms the directionality of the link between risk and α wave amplitude as demonstrated in the previous literature.

# 7 Discussion

The aim of the current paper was to jointly model behavioural and neural data in the hybrid choice model framework to gain insights into cycling behaviour and associated neural processing and consequently, deepen our understanding achieved with the previous models. Therefore, we proposed three hypotheses with respect to observed behaviour and the α amplitude.

The results of the developed hybrid model are in line with our hypotheses. The model estimates indicate that the cyclists in the virtual scenarios are indeed more likely to reduce their speed when approaching the junction where passing cars are present (Hypothesis 1a). They also demonstrate that the reaction to the passing cars becomes stronger as the cyclists get closer to the potential collision (Hypothesis 1b). Further, the model allowed us to quantify the relative impact of a range of influencing variables (such as a cyclist's current speed or distance relative to the junction) on the cyclist's responses. Together, they build a complex picture of cyclist's behaviour where we observe that among them, the most impactful on the choice of current action is the possibility of collision with other road users, providing evidence that our experiment was able to elicit realistic reactions to the traffic events.

Further, the employment of the instrumented bicycle allowed us to capture the physical effort exerted by the cyclists, where we observe that high speed levels increase the physical tiredness and results in a gradual loss of speed. We could also observe that typical speed level developed by the cyclists in our experiment is in line with that observed in other studies and that cyclists naturally converged to that desired level by increasing or decreasing acceleration rate accordingly. In particular, cyclists were more inclined to accelerate or maintain speed if their current speed was below or within the range of average speed developed by the cyclists in reality (Schleinitz et al., 2017).

The inclusion of the continuous element showed the magnitudes of the performed actions allowing us to reach additional level of detail with respect to how cyclists behave. Importantly, the observed results are in line with a priori expectations, which in turn, provide an additional evidence to support the validity for the adopted approach, that is essential in case of exploratory research such as the current one to increase the level of confidence in the findings and methodology. Nonetheless, even though what we observe is plausible from the behavioural perspective, the external validity of current results is yet to be established. Therefore, future studies should aim at the comparison of the magnitudes of cycling actions between virtual reality and the real world to provide evidence of relative and absolute validity of cycling behaviour (Branzi et al., 2017; Hussain et al., 2019). The experimental design encompassing real cycling activity and its replication within the simulated environment, as proposed in the recent study by Guo et al. (2021), would allow for evaluating the reliability of VR experiment in cycling context.

Furthermore, the latent component in the model showed that the decreased α amplitude is associated with the increased perceived risk and elevated complexity of the task faced by the cyclist which in turn increases the propensity to brake which confirmed our hypothesis (Hypothesis 2). The benefits of incorporating the neural data in explaining the behaviour are reflected in the improvement in the efficiency (i.e. smaller standard errors) of variables included in the structural equation (in comparison with a model that does not use the neural data). These results, taken together, demonstrate a behavioural and neural congruence. In this sense, the neural data could be seen as an alternative or at least an addition to attitudinal scales, frequently used in the latent constructs. As these are claimed to be prone to bias of the respondent (conscious or unconscious), heavily dependent on the choice of scale or be

susceptible to the experimenter effect. Therefore, neural data, seen as an unfiltered response, could help offsetting these effects and provide a validation tool if used jointly in the models. The results also demonstrate the feasibility of successful incorporation of neural data into mathematical models in general through the hybrid structure.

Nonetheless, it is crucial to be aware of the potential constraints of this type of work. Even though, in this study we chose α power, which is a relatively well-understood brainwave with a broad spectrum of studies which explored it in different conditions, it is important to take our results with a degree of caution because this is an exploratory work. It attempts to marry difficult disciplines, therefore the promising results presented here are only a small step forward where undoubtedly more research is required to build a strong case. Moreover, the sample composition can be also considered a limitation, where it mainly consisted of students and staff of the National University of Singapore, therefore a future study on a more diverse population and larger sample size would allow for better capturing the role of demographics and improve the generalisability of the findings. Another potential limitation of this work are the aspects in the scenario such as the elements of urban infrastructure or features of other agents that also have influence on cyclists' choices as suggested by the significant alternative specific constants in the model outputs, but which are not accounted for in the model. Future work should thus consider inclusion of various other variables.

Overall, the practical implications of this study are three-fold. Firstly, this study extends the practical knowledge on the VR study design for those researchers who plan to implement it in the future, where we were able to capture the complex behaviour and reactions of the cyclists depending on the changing situation on the road. This, builds and increases the confidence in the validity of VR studies in transport context. Secondly, we provide evidence that behavioural and neural data can be collected jointly using state-of-art equipment currently available on the market with respect to HDMs and neuroimaging. Thirdly, we demonstrate how neural data can be incorporated into the mathematical models and provide an example of an integrative approach to understanding human choices in the dynamic context. This serves as the first step in bridging the gap between mathematical modelling and neuroscience and is expected to encourage further research in this direction.

# Appendices

# Appendix A

In this appendix, we present the full output of the reduced discrete-continuous model, as described in section 6.

*Table 4: A reduced discrete-continuous model (robust standard errors and t-ratios in brackets).*

| |
|---|
| LL(final, whole model): -74840.15 |
| AIC: 149762.3 |
| BIC: 150126.2 |
| LL(final, choice component): -45088.89 |
| LL(final, acceleration when static): 464.95 |
| LL(final, acceleration when moving): -14384.50 |

LL(final, deceleration): -7492.83
LL(final, braking deceleration): -8338.87

| | | Action | Estimate (rob. std. err.; rob.t-ratios) |
|---|---|---|---|
| ASC ($\delta$) | | Maintain | -0.2479 (0.0734; -3.38) |
| | | Brake | -2.1230 (0.0988; -21.49) |
| | | Decelerate | -2.8950 (0.1129; -25.65) |
| | | Wait | 4.1069 (0.1869; 21.98) |
| | | Accelerate | 0 |
| Individual's heterogeneity ($\beta_\sigma$) | | Maintain | 0.2650 (0.0301; 8.80) |
| | | Decelerate | 0.1763 (0.0274; 6.42) |
| Collision | | Maintain | 0.2007 (0.0389; 5.16) |
| | | Brake | 0.7376 (0.1529; 4.82) |
| | | Decelerate | 0.4067 (0.1249; 3.26) |
| | | Wait | 0.6291 (0.1758; 3.58) |
| Time to collision | | Brake | -0.5257 (0.0916; -5.74) |
| | | Decelerate | -0.2144 (0.1459; -1.47) |
| Distance to junction | 1st order | Maintain | 0.0458 (0.0624; -4.70) |
| | 2nd order | Maintain | -0.0835 (0.0079; 2.73) |
| | 1st order | Brake | 0.1410 (0.0184; 7.66) |
| | 2nd order | Brake | -0.2661 (0.0683; -3.90) |
| | 3rd order | Brake | 0.0187 (0.0089; 2.10) |
| | 1st order | Decelerate | 0.0967 (0.0124; 7.82) |
| | 2nd order | Decelerate | -0.2931 (0.0624; -4.70) |
| | 3rd order | Decelerate | 0.0216 (0.0079; 2.73) |
| Speed | 1st order | Maintain | 0.0606 (0.0062; 9.71) |
| | 2nd order | Maintain | -0.0025 (0.0005; -5.61) |
| | 1st order | Brake | 0.0481 (0.0137; 3.51) |
| | 2nd order | Brake | 0.0013 (0.0006; 2.22) |
| | 1st order | Decelerate | 0.1265 (0.0151; 8.40) |
| | 2nd order | Decelerate | -0.0031 (0.0044; -0.70) |
| | 3rd order | Decelerate | 0.0028 (0.0050; 0.56) |
| Continuous acceleration when stopped | | μ | -3.0631 (0.1386; -22.10) |
| | | σ | 2.2931 (0.0795; 28.84) |
| Continuous acceleration when in motion | | μ | -0.3722 (0.0336; -11.09) |
| | | σ | 1.0912 (0.0144; 75.58) |
| | | Speed | 0.0518 (0.0022; 23.90) |
| Continuous deceleration | | μ | -0.2864 (0.0435; -6.59) |
| | | σ | 1.1047 (0.0196; 56.38) |
| | | Speed | 0.0394 (0.0026; 15.15) |
| Continuous braking deceleration | | μ | -0.4394 (0.0639; -6.87) |
| | | σ | 1.1958 (0.0253; 47.32) |
| | | Speed | 0.0602 (0.0038; 15.97) |
| Latent risk (θ) | | Brake ($\beta_{risk_{Br}}$) | 0.1159 (0.0176; 6.58) |
| | | Distance to the junction ($\gamma_{dist-to-junction}$) | -0.0524 (0.0305; -1.72) |
| | | Car at the junction ($\gamma_{car-present}$) | 0.0047 (0.0021; 2.27) |
| | | ζα | - |

# Appendix B

*Table 5: A reduced model without the continuous component (robust standard errors and t-ratios in brackets).*

LL (final, whole model): -119891.1
AIC: 239846.20
BIC: 240130.20

LL (final, MNL): -45129.09
LL (final α): -74790.59

| | | Action | Estimate (rob.t-ratios) |
|---|---|---|---|
| ASC ($\delta$) | | Maintain | -0.2644 (0.0797; -3.32) |
| | | Brake | -2.4598 (0.1256; -19.58) |
| | | Decelerate | -2.9847 (0.1183; -25.23) |
| | | Wait | 4.1069 (0.1869; 21.98) |
| | | Accelerate | 0 |
| Individual's heterogeneity ($\beta_\sigma$) | | Maintain | 0.2241 (0.0217; 10.34) |
| | | Brake | 0.0363 (0.0397; 0.92) |
| | | Decelerate | -0.1051 (0.0445; -2.36) |
| Collision | | Maintain | 0.2031 (0.0384; 5.28) |
| | | Brake | 0.7535 (0.1542; 4.89) |
| | | Decelerate | 0.3616 (0.1185; 3.05) |
| | | Wait | 0.6292 (0.1758; 3.58) |
| Time to collision | | Brake | -0.5280 (0.0905; -5.83) |
| | | Decelerate | -0.1612 (0.1508; -1.07) |
| Distance to junction | 1st order | Maintain | 0.0444 (0.0038; 11.68) |
| | 2nd order | Maintain | -0.0835 (0.0085; -9.86) |
| | 1st order | Brake | 0.1900 (0.0138; 13.78) |
| | 2nd order | Brake | -0.2708 (0.0604; -4.48) |
| | 3rd order | Brake | 0.0195 (0.0080; 2.44) |
| | 1st order | Decelerate | 0.0850 (0.0125; 6.78) |
| | 2nd order | Decelerate | -0.2446 (0.0621; -3.94) |
| | 3rd order | Decelerate | 0.0166 (0.0078; 2.12) |
| Speed | 1st order | Maintain | 0.0605 (0.0062; 9.80) |
| | 2nd order | Maintain | -0.0026 (0.0004; -5.79) |
| | 1st order | Brake | 0.0458 (0.0141; 3.26) |
| | 2nd order | Brake | 0.0014 (0.0006; 2.25) |
| | 1st order | Decelerate | 0.1931 (0.0270; 7.16) |
| | 2nd order | Decelerate | -0.0058 (0.0019; -3.09) |
| | 3rd order | Decelerate | 0.0001 (0.0000; 3.06) |
| Latent risk (θ) | | Brake ($\beta_{risk_{Br}}$) | 0.3285 (0.0774; 4.24) |
| | | Distance to the junction ($\gamma_{dist-to-junction}$) | -0.0762 (0.0058; -13.11) |
| | | Car at the junction ($\gamma_{car-present}$) | 0.0041 (0.0015; 2.72) |
| | | ζα | -0.0937 (0.0187; -5.01) |


## Acknowledgements

Martyna Bogacz, Stephane Hess, Charisma Choudhury and Chiara Calastri acknowledge the support of the European Research Council through the consolidator grant 615596-DECISIONS.


## References


Abou-Zeid, M. and Ben-Akiva, M. (2014). Hybrid choice models. *In Handbook of choice modelling*. Edward Elgar Publishing.

Ambekar, A. and Achrekar, V. (2014). Real time EGG measurement. Delta, 1(5):20–200.
Andersson, H. (2013). Consistency in preferences for road safety: An analysis of precautionary and stated behavior. *Research in transportation economics*, 43(1):41–49.

Awais, M., Badruddin, N., and Drieberg, M. (2017). A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. *Sensors*, 17(9):1991.

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

Bennett, R. and Vijaygopal, R. (2018). Consumer attitudes towards electric vehicles. *European Journal of Marketing*, 53(3/4):499–527.

Berg, P. and Scherg, M. (1994). A multiple source approach to the correction of eye artifacts. *Electroencephalography and clinical neurophysiology*, 90(3):229–241.

Bogacz, M., Hess, S., Calastri, C., Choudhury, C. F., Erath, A., van Eggermond, M. A., Mushtaq, F., Nazemi, M., and Awais, M. (2020). Comparison of cycling behavior between keyboard-controlled and instrumented bicycle experiments in virtual reality. *Transportation research record*, 2674(7):244–257.

Bolduc, D. and Alvarez-Daziano, R. (2010). On estimation of hybrid choice models. In Choice Modelling: The State-of-the-Art and the State-of-Practice: *Proceedings from the Inaugural International Choice Modelling Conference. Emerald Group Publishing*, page 259.

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, 44:58–75.

Borrego, A., Latorre, J., Alcaniz, M., and Llorens, R. (2018). Comparison of oculus rift and htc vive: feasibility for virtual reality-based exploration, navigation, exergaming, and rehabilitation. *Games for health journal*, 7(3):151–156.

Branzi, V., Domenichini, L., and La Torre, F. (2017). Drivers' speed behaviour in real and simulated urban roads–a validation study. *Transportation research part F: traffic psychology and behaviour*, 49:1–17.

Brouwer, A.-M., Neerincx, M. A., Kallen, V., van der Leer, L., and ten Brinke, M. (2011). EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress. *Journal of Cybertherapy & Rehabilitation*, 4(1):21–34.

Camara, F., Dickinson, P., and Fox, C. (2021). Evaluating pedestrian interaction pref- erences with a game theoretic autonomous vehicle in virtual reality. *Transportation research part F: traffic psychology and behaviour*, 78:410–423.

Chirico, A., Cipresso, P., Yaden, D. B., Biassoni, F., Riva, G., and Gaggioli, A. (2017). Effectiveness of immersive videos in inducing awe: an experimental study. *Scientific Reports*, 7(1):1–11.

Christiano, L. J. and Fitzgerald, T. J. (2003). The band pass filter. international economic review, 44(2):435–465.

Cooper, N. R., Burgess, A. P., Croft, R. J., and Gruzelier, J. H. (2006). Investigat- ing evoked and induced electroencephalogram activity in task-related alpha power increases during an internally directed attention task. *Neuroreport*, 17(2):205–208.

De Houwer, J., Beckers, T., and Moors, A. (2007). Novel attitudes can be faked on the implicit association test. *Journal of Experimental Social Psychology*, 43(6):972–978.

Dixit, V. V., Harb, R. C., Mart´ınez-Correa, J., and Rutstr¨om, E. E. (2015). Measuring risk aversion to guide transportation policy: Contexts, incentives, and respondents. *Transportation Research Part A: Policy and Practice*, 80:15–34.

Fairclough, S. H., Venables, L., and Tattersall, A. (2005). The influence of task de- mand and learning on the psychophysiological response. *International Journal of Psychophysiology*, 56(2):171–184.

Fournier, L. R., Wilson, G. F., and Swain, C. R. (1999). Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training. *International Journal of Psychophysiology*, 31(2):129–145.

Frankenhuis, W. E., Dotsch, R., Karremans, J. C., and Wigboldus, D. H. (2010). Male physical risk taking in a virtual environment. *Journal of Evolutionary Psychology*, 8(1):75–86.

Furman, G. D. and Baharav, A. (2010). Investigation of drowsiness while driving utilizing analysis of heart rate fluctuations. *IEEE*.

Guo, X., Robartes, E. M., Angulo, A., Chen, T. D., and Heydarian, A. (2021). Bench- marking the use of immersive virtual bike simulators for understanding cyclist behaviors.

Haghani, M., Bliemer, M. C., Farooq, B., Kim, I., Li, Z., Oh, C., Shahhoseini, Z., and MacDougall, H. (2021). Applications of brain imaging methods in driving behaviour research. *Accident Analysis & Prevention*, 154:106093.

Haufe, S., Treder, M. S., Gugler, M. F., Sagebaum, M., Curio, G., and Blankertz, B. (2011). Eeg potentials predict upcoming emergency brakings during simulated driving. *Journal of neural engineering*, 8(5):056001.

Hern´andez, L. G., Mozos, O. M., Ferr´andez, J. M., and Antelis, J. M. (2018). Eeg- based detection of braking intention under different car driving conditions. *Frontiers in neuroinformatics*, 12:29.

Hess, S. and Palma, D. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of choice modelling*, 32:100170.

Hu, S., Bowlds, R. L., Gu, Y., and Yu, X. (2009). Pulse wave sensor for non-intrusive driver's drowsiness detection. *In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2312–2315. IEEE.

Hussain, Q., Alhajyaseen, W. K., Pirdavani, A., Reinolsmann, N., Brijs, K., and Brijs,T. (2019). Speed perception and actual speed in a driving simulator and real-world: A validation study. *Transportation research part F: traffic psychology and behaviour*, 62:637–650.

Ille, N., Berg, P., and Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of clinical neurophysiology*, 19(2):113–124.

Isa, I. S., Zainuddin, B. S., Hussain, Z., and Sulaiman, S. N. (2014). Preliminary study on analyzing EEG alpha brainwave signal activities based on visual stimulation. *Procedia Computer Science*, 42:85–92.

Iyengar, S. (2011). Laboratory experiments in political science. In Druckman, J. N., Greene, D. P., Kuklinski, J. H., and Lupia, A., editors, *Cambridge handbook of experimental political science*, pages 73–88. Cambridge University Press, Cambridge, Cambridge.

Johnson, M., Charlton, J., Oxley, J., and Newstead, S. (2010). Naturalistic cycling study: identifying risk factors for on-road commuter cyclists. *In Annals of advances in automotive medicine/annual scientific conference*, volume 54, page 275. Association for the Advancement of Automotive Medicine.

Johnson, M. J., Chahal, T., Stinchcombe, A., Mullen, N., Weaver, B., and Bedard, M. (2011). Physiological responses to simulated and on-road driving. *International Journal of Psychophysiology*, 81(3):203-208.

Kim, D.-Y. (2003). Voluntary controllability of the implicit association test (IAT). *Social Psychology Quarterly*, pages 83–96.

Kim, I.-H., Kim, J.-W., Haufe, S., and Lee, S.-W. (2014). Detection of braking intention in diverse situations during simulated driving based on EEG feature combination. *Journal of neural engineering*, 12(1):016001.

Koutsopoulos, H. N. and Farah, H. (2012). Latent class model for car following behavior. *Transportation research part B: methodological*, 46(5):563–578.

Lei, S. and Roetting, M. (2011). Influence of task combination on EEG spectrum modulation for driver workload estimation. *Human factors*, 53(2):168–179.

Lewis, R. S., Weekes, N. Y., and Wang, T. H. (2007). The effect of a naturalistic stressor on frontal EEG asymmetry, stress, and health. *Biological psychology*, 75(3):239–247.

Magosso, E., De Crescenzio, F., Ricci, G., Piastra, S., and Ursino, M. (2019). EEG alpha power is modulated by attentional changes during cognitive tasks and virtual reality immersion. *Computational intelligence and neuroscience*, 2019.

Mai, K. L. (2017). Evaluation of pc-based virtual reality as a tool to analyze pedestrian behavior at midblock crossings.

Makeig, S. and Jung, T.-P. (1995). Changes in alertness are a principal component of variance in the eeg spectrum. *NeuroReport-International Journal for Rapid Communications of Research in Neuroscience*, 7(1):213–216.

Malfetti, J. L. et al. (1989). *Young Driver Attitude Scale: The Development and Field-Testing of an Instrument To Measure Young Driver Risk-Taking Attitudes*. ERIC, Columbia University, New York, NY. Teachers College.

Mann, C. A., Sterman, M. B., and Kaiser, D. A. (1996). Suppression of EEG rhythmic frequencies during somato-motor and visuo-motor behavior. *International Journal of Psychophysiology*, 23(1-2):1–7.

Michon, J. A. (1985). A critical view of driver behavior models: what do we know, what should we do? *In Human behavior and traffic safety*, pages 485–524. Springer, Boston, MA.

Mummolo, J. and Peterson, E. (2019). Demand effects in survey experiments: An empirical assessment. *American Political Science Review*, 113(2):517–529.

Näher, A.-F. and Krumpal, I. (2012). Asking sensitive questions: the impact of forgiving wording and question context on social desirability bias. *Quality & Quantity*, 46(5):1601–1616.

Nishifuji, S., Sato, M., Maino, D., and Tanaka, S. (2010). Effect of acoustic stimuli and mental task on alpha, beta and gamma rhythms in brain wave. In *Proceedings of SICE Annual Conference 2010, pages 1548–1554.* IEEE.

Nuñez Velasco, J. P., de Vries, A., Farah, H., van Arem, B., and Hagenzieker, M. P. (2021). Cyclists' crossing intentions when interacting with automated vehicles: A virtual reality study. *Information*, 12(1):7.

Paschalidis, E., Choudhury, C. F., and 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.

Patterson, P. G. and Mattila, A. S. (2008). An examination of the impact of cultural orientation and familiarity in service encounter evaluations. *International Journal of Service Industry Management*, 19:662–681.

PTV-Group (2021). *Ptv vissim new*.

Ram, T. and Chand, K. (2016). Effect of drivers' risk perception and perception of driving tasks on road safety attitude. *Transportation research part F: traffic psychology and behaviour*, 42:162–176.

Ray, W. J. and Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*, 228(4700): 750–752.

Reason, J., Manstead, A., Stradling, S., Baxter, J., and Campbell, K. (1990). Errors and violations on the roads: a real distinction? *Ergonomics*, 33(10-11):1315–1332.

Riera, A., Dunne, S., Cester, I., and Ruffini, G. (2008). Starfast: A wireless wearable EEG/ECG biometric system based on the Enobio sensor. In *Proceedings of the international workshop on wearable micro and nanosystems for personalised health*.

Román, C., Martín, J. C., and Espino, R. (2014). Using stated preferences to analyze the service quality of public transport. *International Journal of Sustainable Transportation*, 8(1):28–46.

Rundmo, T. and Iversen, H. (2004). Risk perception and driving behaviour among adolescents in two norwegian counties before and after a traffic safety campaign. *Safety science*, 42(1):1–21.

Schleinitz, K., Petzoldt, T., Franke-Bartholdt, L., Krems, J., and Gehlert, T. (2017). The german naturalistic cycling study–comparing cycling speed of riders of different e-bikes and conventional bicycles. *Safety Science*, 92:290–297.

Seo, S.-H. and Lee, J.-T. (2010). Stress and EEG. *Convergence and hybrid information technologies*, 1(1):413–424.

Shechtman, O., Classen, S., Awadzi, K., and Mann, W. (2009). Comparison of driving errors between on-the-road and simulated driving assessment: a validation study. *Traffic injury prevention*, 10(4):379–385.

Simon, A. J., Schachtner, J. N., and Gallen, C. L. (2019). Disentangling expectation from selective attention during perceptual decision making. *Journal of neurophysiology*, 121(6):1977–1980.

Slobounov, S., Fukada, K., Simon, R., Rearick, M., and Ray, W. (2000). Neurophysiological and behavioral indices of time pressure effects on visuomotor task performance. *Cognitive Brain Research*, 9(3):287–298.

Svensson, M. (2009). Precautionary behavior and willingness to pay for a mortality risk reduction: Searching for the expected relationship. Journal of risk and uncertainty, 39(1):65–85.

Teplan, M. et al. (2002). Fundamentals of EEG measurement. *Measurement science review*, 2(2):1–11.

Ulleberg, P. and Rundmo, T. (2003). Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. *Safety science*, 41(5):427– 443.

Underwood, G., Crundall, D., and Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation research part F: traffic psychology and behaviour*, 14(6):435–446.

Unity (2017). *Unity - game engine*.

Vanitha, V. and Krishnan, P. (2016). Real time stress detection system based on EEG signals.

Vanni, S., Revonsuo, A., and Hari, R. (1997). Modulation of the parieto-occipital alpha rhythm during object detection. *Journal of Neuroscience*, 17(18):7141–7147.

Varotto, S., Farah, H., Toledo, T., van Arem, B., and Hoogendoorn, S. (2018). Continuous-discrete choices of control transitions and speed regulations in full-range adaptive cruise control. In *TRB 2018: 97th Annual Meeting of the Transportation Research Board. Transportation Research Board (TRB).*

Verona, E., Sadeh, N., and Curtin, J. J. (2009). Stress-induced asymmetric frontal brain activity and aggression risk. *Journal of abnormal psychology*, 118(1):131.

Welch, P. (1967). The use of fast fourier transform for the estimation of power spec- tra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73.

Zhang, Q., Fu, R., Guo, Y., Guo, Y., Yuan, W., Wang, C., Wu, F., and Ma, Y. (2011). Risk attitude, perception, behavior, and personality as indicators of a driver's risk awareness in China. In *3rd International Conference on Road Safety and Simulation*, pages 1–13.