A Comparison of Cycling Behaviour between Keyboard-Controlled and Instrumented Bicycle Experiments in Virtual Reality

Martyna Bogacz
Institute for Transport Studies & Choice Modelling Centre
University of Leeds
34-40 University Road, Leeds LS2 9JT
Email: tsmb@leeds.ac.uk

Stephane Hess
Institute for Transport Studies & Choice Modelling Centre
University of Leeds
34-40 University Road, Leeds LS2 9JT
Email: s.hess@leeds.ac.uk

Chiara Calastri
Institute for Transport Studies & Choice Modelling Centre
University of Leeds
34-40 University Road, Leeds LS2 9JT
Email: c.calastri@leeds.ac.uk

Charisma Choudhury
Institute for Transport Studies and & Choice Modelling Centre
University of Leeds
34-40 University Road, Leeds LS2 9JT
Email: c.f.choudhury@leeds.ac.uk

Alex Erath
Singapore-ETH Centre
1 Create Way, CREATE Tower, #06-01, Singapore 138602
Email: erath@erveco.ch

Michael van Eggermond
The University of Applied Sciences and Arts of Western Switzerland
Hofackerstrasse 30, 4132 Muttenz, Switzerland
Email: michael.vaneggermond@fhnw.ch

Faisal Mushtaq
Institute of Psychological Sciences
University of Leeds
4 Lifton Pl, Leeds LS2 9JZ
Email: f.mushtaq@leeds.ac.uk

Mohsen Nazemi
Singapore-ETH Centre
1 Create Way, CREATE Tower, #06-01, Singapore 138602
Email: nazemi@ivt.baug.ethz.ch

Muhammad Awais
Energy and Environmental Institute

Bogacz, M., Hess, S., Calastri, C., Choudhury, C. F., Erath, A., van Eggermond, M. A., Mushtaq, F., Nazemi, M. & Awais, M.

University of Hull
Hull, HU6 7RX
Email: mawais102@gmail.com

Word Count: 8083 words + 2 table (250 words per table) = 8583 words
Submitted [26/03/2020]

Bogacz, M., Hess, S., Calastri, C., Choudhury, C. F., Erath, A., van Eggermond, M. A., Mushtaq, F., Nazemi, M. & Awais, M.

ABSTRACT

The use of virtual reality (VR) in transport research offers the opportunity to collect behavioural data in a controlled dynamic setting. VR settings are useful in the context of hypothetical situations where real-world data does not exist and/or in situations which involve risk and safety issues making real-world data collection infeasible. Nevertheless, VR studies can contribute to transport-related research only if the behaviour elicited in a virtual environment closely resembles real-world behaviour. Importantly, as VR is a relatively new research tool, the best-practice in terms of the experimental design is still to be established. In this paper, we contribute to a better understanding of the implications of the choice of the experimental setup by comparing cycling behaviour in VR between two groups of participants in similar immersive scenarios – the first group controlling the manoeuvres using a keyboard and the other group riding an instrumented bicycle. We critically compare the speed, acceleration, braking and head movements of the participants in the two experiments. We also collect electroencephalography (EEG) data to compare the alpha wave amplitudes and assess the engagement levels of participants in the two settings. The results demonstrate the ability of VR to elicit behavioural patterns in line with those observed in the real-world and indicate the importance of the experimental design in a VR environment beyond the choice of audio-visual stimuli. The findings will be useful for researchers in designing the experimental setup of the VR for behavioural data collection.

Keywords: virtual reality, instrumented bicycle, keyboard, EEG

INTRODUCTION

Virtual reality (VR) has become an increasingly popular tool for travel behaviour research. This is because in the transport domain, it is often inherently difficult to collect real-life data in hazardous road circumstances. VR provides a high degree of experimental control, safety and ease of data collection, while at the same time allowing to collect data in a dynamic setting. Further, as in other domains, VR makes it possible to collect data in hypothetical future scenarios allowing to pre-test behavioural responses in the context of new modes and novel urban design. Consequently, it has been widely used in previous studies in a transport context. For example, Mai (1) evaluated VR as a tool to analyse pedestrian behaviour at midblock crossings and Frankenhuis et al. (2) explored male risk-taking behaviour while crossing a bridge in an immersive environment. Finally, Moussa et al. (3) tried to apply the Augmented Reality Vehicle system to left-turn manoeuvres at two-way stop-controlled intersections.

Nevertheless, the potential disadvantages of VR include motion sickness, high costs and most importantly, the risk of an unrealistic representation of reality. The ecological validity of VR experiments is one of its main issues, as it is widely known that experimental designs which have a different degree of immersion or employ different equipment can elicit distinct behavioural responses. For example, Farooq et al. (4) elicited preferences over Connected and Autonomous Vehicles comparing three methods: an immersive reality technology, a conventional visual presentation and text-only descriptions. The findings showed that preference for autonomous vehicles increased from 40% in text-only case, through 50% in a visual presentation to 70% if VR was used. It was concluded that preferences elicited with immersive equipment were more consistent with real world preferences and the understanding of scenarios improved. Furthermore, Bogacz et al. (5), which looked at the differences in risk processing between 2D and 3D cycling scenarios in VR, showed that the behavioural patterns from the experiments were similar to the actual behaviour of cyclists on the roads. Moreover, the study found that the propensity to brake was higher in the 3D presentations compared to the 2D scenarios. Patterson et al. (6) conducted two experiments to investigate the influence of the presentation method on neighbourhood choice. The first one was based on the textual description of the living area, while the second used VR simulations of the neighbourhood. The results showed that preferences elicited with text-only surveys reflected participants' subjective, imagined illustration of the described place, whereas, in the case of the visualizations, the preferences were based on the observed material. It suggests that VR technology allows for constructing experimental scenarios that give the researcher more control over factors that affect the respondents' choices, but at the same time these studies clearly exhibit the fact that preferences are highly dependent on the presentation format. This effect is expected to be even stronger within immersive technology experiments, as they engage individuals to a larger extent than traditional survey methods. However, as VR is still an innovative and relatively new research tool, the best practices are yet to be established especially in the light of mixed evidence in the existing literature.

In contrast, the domain of driving simulators represents an exception, as there has been extensive research on the factors that affect the behavioural congruence in simulated driving. For instance, Underwood et al. (7), who assessed the comparability of driving on a road and in a simulator, concluded that driving simulators can demonstrate similar patterns of differences across drivers as observed on actual roads. However, this was only relative, in the sense that they were unable to create the same hazardous situations on a road as can be designed in a simulator. Furthermore, Godley (8) examined the validity of driving simulators by comparing driving behaviour in an instrumented car vs a simulator. They showed similar deceleration activity under both conditions. Yet, on the other hand, individuals tended to drive faster in the instrumented car relative to the simulator. To the best of our knowledge, there has not been any similar in-depth investigation on the factors that affect the behavioural congruence in the context of different variants of immersive cycling environments.

In this paper, we address this research gap and contribute to a better understanding of the implications of the choice of the experimental setup by comparing the cycling behaviour in VR between two groups of participants in similar immersive scenarios – the first group controlling the manoeuvres using a keyboard and the other group riding an instrumented bicycle.

These two types of equipment both have their advantages and drawbacks. The use of a keyboard significantly reduces the cost of the experiment as well as the setup time but diminishes the realism of the experiment. The employment of an instrumented bicycle in the experiment is more effort consuming and requires novel engineering design (e.g. measuring wheel and pedal rotation, braking force etc.) and is not portable. Nevertheless, the latter provides participants with an experience which resembles reality to a larger extent and allows analysts to obtain richer data from the sensors installed on the bicycle.

This is also likely to be reflected in their neural activity, which can be used as an indicator of the level of engagement.

There are several studies in cognitive psychology which have looked at the effects of the use of different input devices on the neural processing and the elicited behaviours. For example, key presses can be considered discrete decisions and while they have been widely used in PC-based experiments before (9), many studies have shown a continuous flow of information between the brain systems involved in motor processes as opposed to previous assumptions about sequential stages in motor outputs (10-11). These findings suggest that devices which allow for continuous rather than discrete input in terms of motor decisions better mimic the neural processing of such decisions. Moreover, a study by Rupp et al. (12) demonstrated that the use of a joystick, as opposed to a keyboard, resulted in lower mental workload in a difficult task, which could suggest that keyboards are unsuitable input devices in complex control tasks. Finally, a study by Chung et al. (13) investigated online shopping experience and purchase patterns using both mouse-controlled and touch interface settings. They found that shoppers who used a touch interface to browse products (vs. mouse) have a significantly higher engagement with their shopping experience. The studies mentioned above show that there are possible differences in behaviour resulting from the type of input device adopted, making the search for and testing of alternative solutions in dynamic experiments a research priority.

In addition to comparing cycling behaviour in VR when using different devices to elicit preferences, we also set out to explore the latter's impact on participants' neural activity as a proxy measure of engagement. For this reason, we employed electroencephalography (EEG), a scalp-recorded measure of the electrical activity generated by the brain. Typically, in the transport literature, the use of EEG has largely focused on the investigation of driver fatigue and drowsiness (14-17), level of alertness/attention or cognitive performance (18). However, little has been done to evaluate the engagement of participants in the immersive environment from a neural perspective. In particular, the use of neuroimaging devices in applied experimental research has been heavily constrained by the signal-to-noise ratio of EEG, where artefacts in the data can stem from physiological (e.g. ocular, facial and body muscle movements) and non-physiological sources (e.g. electric signals generated by nearby equipment, as shown by Puce et al. (19)). Therefore, VR experiments which allow a great degree of flexibility in participants' head and body movements are more prone to producing artefactual data. However, recent wireless systems such as Emotiv EPOC+ (20) and Enobio (21) are designed for dynamic experimental setups and attempt to mitigate the impact of movement artefacts on the scalp-recorded EEG.

In this paper, we compare a particular pattern of oscillatory brain activity known as occipital alpha (α) to infer participant's engagement in the task. Occipital α, which is quantified through frequency analysis of the signal ranging from 8 to14 Hz, is one of the most commonly observed signatures of brain activity, with numerous studies demonstrating a relationship between oscillations in this frequency band and attentional processing (22-25). As such, the signal presents an ideal candidate to investigate the impact of presentation format on participants' degree of task-relevant engagement.

The remainder of this paper is organized as follows. We present our specific hypotheses in the next section. The survey design and sample characteristics for the two experiments are discussed next, followed by the methodological approach of the study. We next turn to the results section, followed by the discussion that reviews the insights from the analysis.

HYPOTHESES

Five hypotheses are put forward based on the evidence from the existing literature presented above and tested empirically using our data. They relate to cycling speed, head movements (an indicator of engagement with the surroundings beyond peripheral vision), acceleration and braking behaviour as well as neural processing. We now look at these five in turn.

Cycling speed

Hypothesis 1A: The average speed is higher in the keyboard-controlled experiment as opposed to instrumented bicycle one.

Hypothesis 1B: There is more variance in speed in the instrumented bicycle experiment than in the keyboard experiment.

It is hypothesized that the average speed will be higher in the keyboard-controlled experiment as it requires less physical effort to accelerate compared to an instrumented bicycle. Moreover, the acceleration is more instantaneous when using the keyboard. For the same reason, we expect that there will be more variation in the speed observed in the instrumented bicycle experiment, as more physical effort is needed to move through the scenarios, making it more difficult to maintain constant speed levels.

Head movement

Hypothesis 2: The average head movement is higher in the instrumented bicycle experiment than in the keyboard experiment.

We expect that the use of the instrumented bicycle will induce participants to inspect the environment, resulting in more head movement (26-27). This would be due to the higher level of immersion in the environment, due to the improved design compared to the keyboard, and due to the fact that braking in case of any hazardous circumstances on the road will take longer on the instrumented bike compared to instantaneous reaction while pressing the arrows on the keyboard.

Acceleration & Braking

Hypothesis 3: There is more variance in the acceleration behaviour in the instrumented bicycle experiment than in the keyboard experiment.

Hypothesis 4: There is more variance in the braking behaviour in the instrumented bicycle experiment than in the keyboard experiment.

Hypotheses 3 and 4 stem from the fact that the use of the instrumented bicycle provides more scope to control behaviour, as participants intertwine acceleration and deceleration more often compared to a keyboard. Cycling on the bike requires more physical effort and time to switch between subsequent actions or respond to changing conditions on the road, whereas acceleration and braking are more instantaneous with the keyboard.

Neural processing

Hypothesis 5: The mean amplitude of the alpha wave is higher in keyboard-controlled experiment compared to that of the instrumented bicycle.

Hypothesis 5 is based on the evidence from a large body of work showing the α wave to be a well-established correlate of attentional processing with an increase in amplitude found as participants' attention drifts away from the task (17, 28). On the other hand, current understanding in neuroscience holds that a low α wave implies increased excitability, and thus an increased response to external stimuli

(29-30). Therefore, we hypothesize that if the keyboard-controlled experiment engages participants to a lower extent, the mean α amplitude is expected to be higher as opposed to the instrumented bicycle data.

EXPERIMENTAL DESIGN

This section describes the common experimental procedure used for the two experiments. It also discusses the components that were different between keyboard-controlled and instrumented bike experiment as well as the basic characteristics of the two samples.

Keyboard-controlled experimental setup

The experimental session started with the participant being seated on a chair and having an Emotiv Epoc+ EEG headset (31) and an Oculus Rift VR (32) head mounted display (HMD) placed on their head. The Emotiv headset uses 14 electrodes (at AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4) sampling across the scalp. The system was selected as its compact design allowed it to be used jointly with the VR HMD. As a first step, the baseline brain activity was recorded with the sampling rate of 128 Hz, while participants had their eyes open and focused their gaze on one point on the screen for 15 seconds. The same procedure was then repeated with their eyes closed. Importantly, before the main part of the experiment started, the participants had an opportunity to familiarise themselves with the equipment in a trial run.

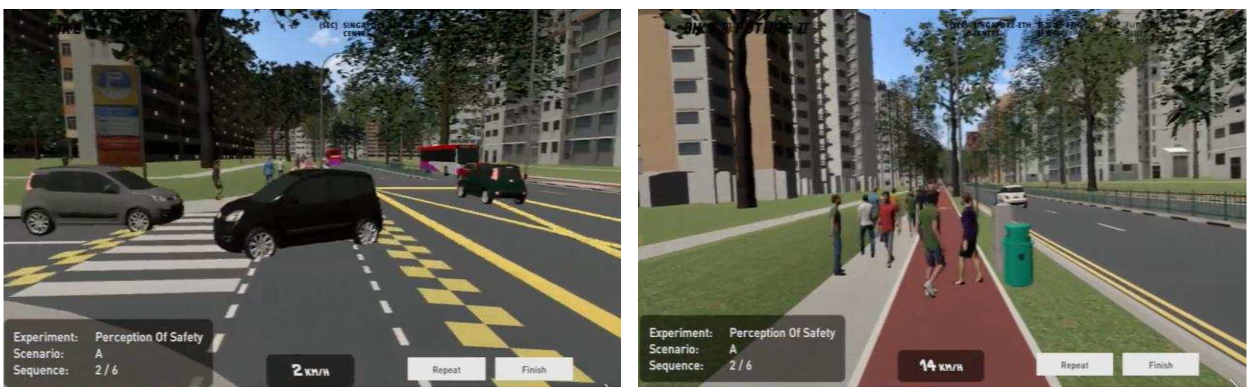

Figure 1: The immersive scenarios used in the experiment.

The experiment used 6 scenarios with an immersive presentation of traffic from the perspective of the cyclist. All the scenarios had a number of common elements. Firstly, the cyclist was riding on the pavement. Secondly, in each scenario, there were four locations where a potential collision with other road users could occur, namely, two junctions and two points along the cycling lane where pedestrians could cross to reach the bin on the right side of the bike lane, as seen in **Figure 1**. Thirdly, all the scenarios featured pedestrians as well as cars and the percentage of pedestrians and cars which would cross the bike lane or turn at the junction was constant. At the same time, there were also random components in each scenario which stemmed from the fact that in each scenario, the specific movements of crossing pedestrians and turning cars were presented randomly, while keeping their overall percentage the same across all the scenarios. This resulted in differences between scenarios in terms of the actual number of pedestrians or cars at the "collision locations" when the cyclist was passing by these points. This was clearly also influenced by the speed of the cyclist, and hence the point in time at which the "collision locations" were reached. Altogether, these elements gave basis for the complex traffic scenarios which participants were required to navigate. The scenarios encompassed a 360-degree view of the road which surrounded the participant and responded to their head movements. Importantly, based on the feedback received during initial pre-testing of the set-up, sound was also included to capture both visual and auditory cues that are available to cyclists in real-life settings. The volume of vehicles was consistent with their distance to the cyclist so that the sound of an approaching car increased as it got closer to the cyclist.

The experiment comprised the same 6 scenarios. The repetition was used because we also collected neural data which requires a higher number of trials in order to obtain more stability in the EEG signal. The task for the participant was to cycle through the scenario at the desired pace until the finish line at the end of each scenario. In order to navigate through the scenario, participants used the keyboard to adjust their speed but had no ability to turn left or right. They pressed the up arrow to accelerate and the down arrow to brake. The keyboard was placed on the table in front of them, and before the experiment began, they were guided by the experimenter to find the appropriate keys on the keyboard. The experimental setup of a keyboard-controlled experiment can be seen on right-hand side of **Figure 2**.

The visual stimuli in the experiment come from VR road simulations developed by Future Cities Laboratory (33) using Unity 3D Game Engine (34). These stimuli involve pre-programmed environments where the cars and pedestrian movements do not respond to the actions of the cyclist. That is to say, other road users do not accelerate or decelerate in reaction to the chosen action of the cyclist (participant), therefore collisions between the cyclist and cars/pedestrians were possible. Collisions were detected if a cyclist overlapped visually with any other agent. Even though participants were specifically instructed to avoid any collisions, there were 19 instances (3.2% of all scenarios by all participants) of collisions with other road users. When this happened, the experiment was interrupted, and the participant was asked to start again from the beginning of that scenario.

The initial number of recruited participants was 50, from which 4 participants were removed due to failure to complete the whole experiment, leading to a final sample size of 46 participants (18 males, 28 females), comprising staff and students of the University of Leeds as well as members of the general public. The mean age of the participants was 30.7 years, with a standard deviation of 10.88 years.

Instrumented bike experimental setup

The experimental design for the instrumented bicycle data collection was similar to the keyboard counterpart in terms of the audio and visual stimuli used and number and types of cycling scenarios. The experimental session commenced by familiarizing the participant with the instrumented bicycle, including demonstrating how to use a hand brake (all the participants could ride a bike in reality). Subsequently, the participant mounted the bike and the HTC Vive head mounted display (HMD) (35) and Enobio (36) devices were placed on their head. The Enobio headset uses 8 electrodes (at FP1, FP2, Fz, C3, Cz, C4, P3 and P4) sampling across the scalp. The system allowed joint use with HTC Vive HMD. As a first step, the baseline brain activity was recorded with the sampling rate of 128 Hz, while participants had their eyes open and focused their gaze on one point on the screen for 1 minute. The same procedure was then repeated with eyes closed. Before the main part of the experiment started, participants could familiarise themselves with the use of the bicycle and the environment in a trial session.

The experiment consisted of 6 immersive presentations of traffic scenarios from the perspective of the cyclists, shown in random order. The visual stimuli proceeded from the same source as before. Similarly, the participant was asked to cycle through the scenario at the desired pace until the finish line at the end of each scenario. In order to navigate through the scenario, the participant used the pedals of the bicycle and had no ability to turn left or right. To brake, the participant used a hand brake positioned on the right side of the handlebar. The instrumented bike can be observed on the left-hand side of **Figure 2**. The instrumented bike belongs to Future Cities Laboratories in Singapore.

Figure 2: Instrumented bicycle (FCL Singapore) and keyboard-controlled experiments (University of Leeds)

The scenarios used in these experiments allowed for recording of cycling behaviour with respect to changing traffic environment. The specific variables of interest, which are described in detail in the next section, included the cycling speed, braking activity, acceleration, horizontal head movements as well as the EEG signal. We put forward several hypotheses based on these variables, as specified in the section above.

Fifty participants were recruited for the experiment, however 2 of them were removed due to failure to complete the whole experiment, leading to a final sample size of 48 participants (29 males, 19 females), comprising staff and students of the National University of Singapore as well as the members of the general public. The mean age of the participants was 26.5 years, with 6.7 years standard deviation.

It is important to emphasize that the small sample size in both experiments is a typical issue faced by researchers working with VR and/or driving simulator data (37- 38, 3) as the experiment duration is much longer and the associated cost much higher compared to typical stated preference (SP) studies.

METHODS

In this section, we present the methodology used to test our hypotheses. We conducted a between-subject comparison of the behaviour in two samples, where one sample used a keyboard and the other used the instrumented bike to cycle through the scenarios. As a result of this experimental design, none of the participants took part in both treatments. To compare the behaviour, we analysed the following variables: acceleration, braking, speed and head movement. We also looked at the difference in mean α amplitude in the two experiments. The sampling rate for all variables except EEG was 4 Hz.

To test the proposed hypotheses, we conduct a Welch's t-test (39) on the mean values of Speed, sideways movements of the head (Head Yaw) as well as α-wave amplitude. This test was chosen due to the slightly unequal sample sizes, where keyboard sample contains 46 participants and the instrumented bicycle sample includes 48 participants. Furthermore, the F-test (40) was used to make inferences about the variances in Speed, Acceleration and Braking between two experiments. The individual variables produced during the experiments were the following:

Acceleration

The acceleration variable (a) is the rate of velocity gain and is measured in metre per squared second (m/s^2). The formula used to calculate acceleration can be seen in **Equation 1**:

$$\mathrm{a} = \frac{\Delta v}{\Delta t} \tag{1}$$

Where Δ denotes changes in velocity (v) and time (t), respectively.

Braking

In the instrumented bicycle experiment, braking is measured as the degree of deviation of the braking pad from its default position. It ranges from 0 to 15^{o}. In the keyboard setting, the braking variable recorded the degree to which the down-arrow key was pressed, and the values range from 0 to 1. In order to be able to compare these values, we performed a min-max normalization on the values of Braking for the instrumented bicycle.

Speed

The speed is expressed in kilometres per hour (km/h). In the keyboard-controlled experiment, the maximum speed was capped at 25 km/h. This level was chosen based on the previous literature which showed that the average speed of cycling in the real world is between 13.5-16 km/h with standard deviation ranging from 3.2 - 8.4 km/h (41-43). Differently, in the instrumented bicycle experiment, there was no limit on the maximum speed. The restriction on the keyboard-controlled experiment was imposed in order to avoid unreasonable speeds, which could have been easily achieved with the constant pressing of the key, and to minimize the risk of motion sickness.

Head movement

The head movement is based on head yaw - the sideways movement of the head. It is measured in degrees from the default position (looking straight ahead), and can range from -180^{o} to $+180^{o}$, where turning the head (as well as the torso) to the left produces negative values while a movement to the right results in positive scores.

EEG

For the EEG analyses, we examined differences in mean α amplitude in keyboard-controlled and instrumented bicycle experiments. As the EEG signal collected through the scalp are inherently noisy, we undertook a number of steps to eliminate artefacts and improve the signal-to-noise ratio. Specifically, we first applied a 1-20 Hz bandpass filtering (BPF), a linear transformation that retains the components of the data within this specific band of frequencies (44) and removes frequencies outside of this range that may stem from physiological sources such as galvanic skin responses or external environmental sources such as electronic equipment (45). Next, we cleaned the data to remove noise stemming from eyeblinks (movement artefacts were corrected using a multiple source analysis method (46-47)). Finally, we computed the power spectrum of the EEG data using Welch's method (48) which estimates the power spectra based on the Fast Fourier Transform (FFT) (49). Because of our interest in occipital α, we performed a region-of-interest analysis and took an average of the activity from electrodes O1, O2, P7, P8, T7 and T8 in the keyboard-experiment and electrodes P7, P8, C3, C4 in the instrumented bicycle to increase the stability of the signal (50).

RESULTS

In this section, we present visual profiles of Speed and Braking for the two experiments and an overview of the descriptive statistics in **Table 1**. The table highlights clear differences between the mean values of the variables of interest in the two samples. Next, these differences are tested more rigorously using t-test and F-test, and the results are reported for each variable.

Table 1: Summary statistics of the variables of interest

Variable	Units of measurement	Mean values in the keyboard-	Standard deviation values in the keyboard-	Mean values in instrumented	Standard deviation values in instrumented

		controlled experiment	**controlled experiment**	**bicycle experiment**	**bicycle experiment**
Speed	km/h	22.90	4.93	14.32	7.03
Head Yaw	degrees	1.76	8.52	4.93	27.71
Break	degrees	0.03	0.10	0.23	0.25
Acceleration	m/s^2	0.11	1.29	0.02	2.68

In **Figure 3**, we plot the profiles of the variables Speed and Braking for two randomly chosen participants cycling in similar immersive scenarios but in the two different experimental settings. The graphs highlight that both experiments captured participants' behaviour correctly, as higher values of Braking are associated with lower Speed. At the same time, it is clear that the employment of the instrumented bike resulted in a considerably higher amount of variation in Speed and Braking compared to the keyboard experiment, where the profiles are smoother and less dynamic.

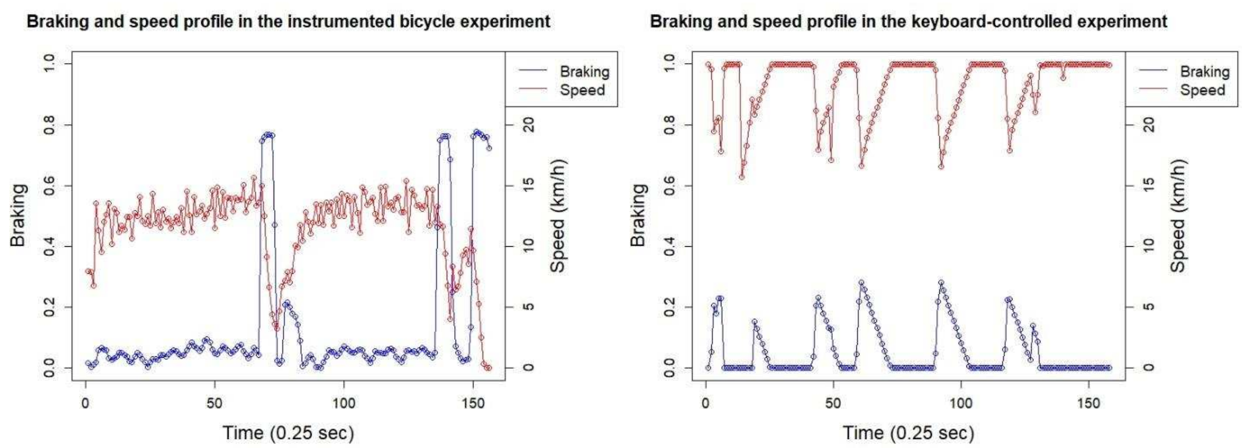

Figure 3: Profiles of Braking and Speed in the two experiments for two randomly chosen participants.

Speed

The one-sided Welch's t-test performed on the mean values of Speed at the individual level showed that the mean Speed is significantly higher in the keyboard-controlled experiment (t = -16.163, df = 70.134, p-value ~ 0). This result is in line with our Hypothesis 1A, where we expected that the lack of physical effort and consequently relative easiness in developing higher speed would result in higher average speed in the keyboard-controlled experiment. Moreover, when we compared the variances in mean Speed, we found that the variance in the instrumented bicycle experiment was significantly higher compared to the keyboard-controlled counterpart (F (47,45) = 3.9252, p-value = 0.000004777). Again, this result conforms to our Hypothesis 1B. We show that the use of the bicycle induces people to adjust their speed more often. It is also interesting to look at the density histogram of Speed in these two experiments (**Figure 4**). We can see the near bell-shaped distribution of Speed in the experiment using the bicycle, where the values are centred around the mean and there is more variation observed in contrast with the keyboard-controlled experiments. A peak near zero can be observed in the instrumented-bicycle setting: this relates to small movements when participants slowed down to stop (while waiting to cross the street or give priority to pedestrians who crossed their bike lane). In contrast, in the keyboard experiment, we observe a skewed distribution of Speed where the majority of observations correspond with the maximum possible level, equal to 25 km/h. This suggests that the removal of physical effort and the use of a keyboard contributes to the choice of maximum speed regardless of the scenario conditions.

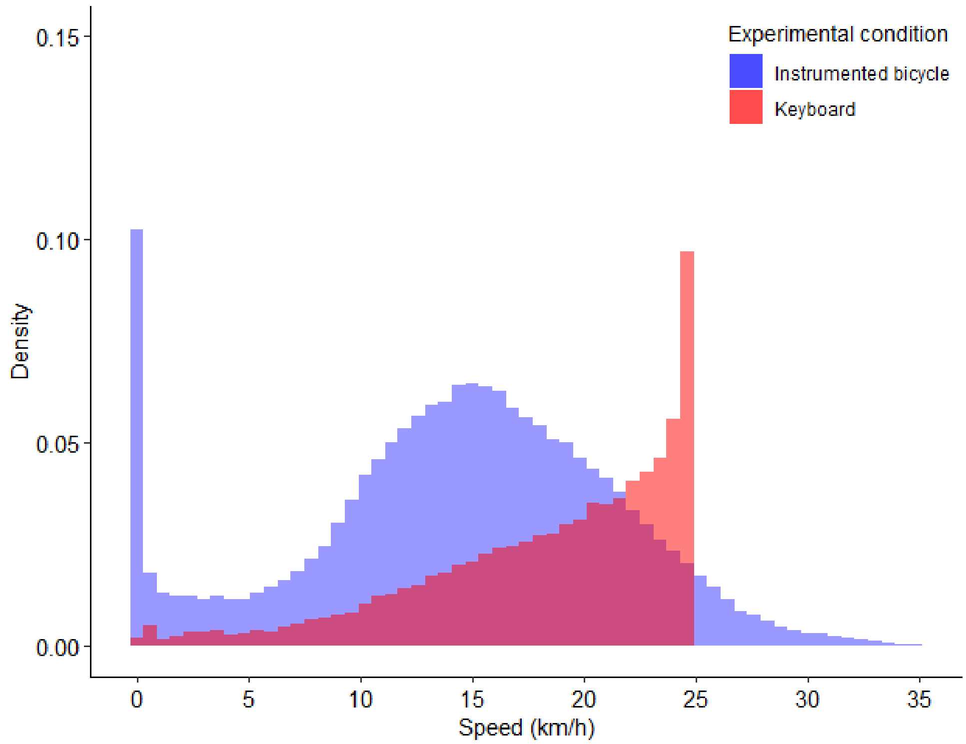

Figure 4: Histograms of Speed in the two experimental settings.

Head movement

Head movement is an indication that the participant is making an effort to gather information beyond the peripheral vision. The one-sided Welch's t-test of Head Yaw demonstrated that the average head movement is significantly higher in the instrumented bicycle experiment as opposed to the keyboard-controlled setting ($t = 2.3944$, $df = 89.987$, p-value = 0.009362). This is in line with our Hypothesis 2, where we expected more head movement to the sides in the instrumented bicycle experiment due to more complex mechanism of control over the bicycle compared to keyboard, which in turn required participants to explore the environment to a larger extent in order to be able to react more quickly. The density histogram of the Head Yaw values (**Figure 5**) shows this trend, as it includes a wider range of values for instrumented bicycle compared to the keyboard experiment. It suggests that the use of the instrumented bicycle induces participants to inspect the environment more than a keyboard due to the higher level of immersion in the environment and due to the fact that braking in case of any hazardous circumstances on the road will take longer on the instrumented bike compared to instantaneous reaction while pressing the arrow on the keyboard.

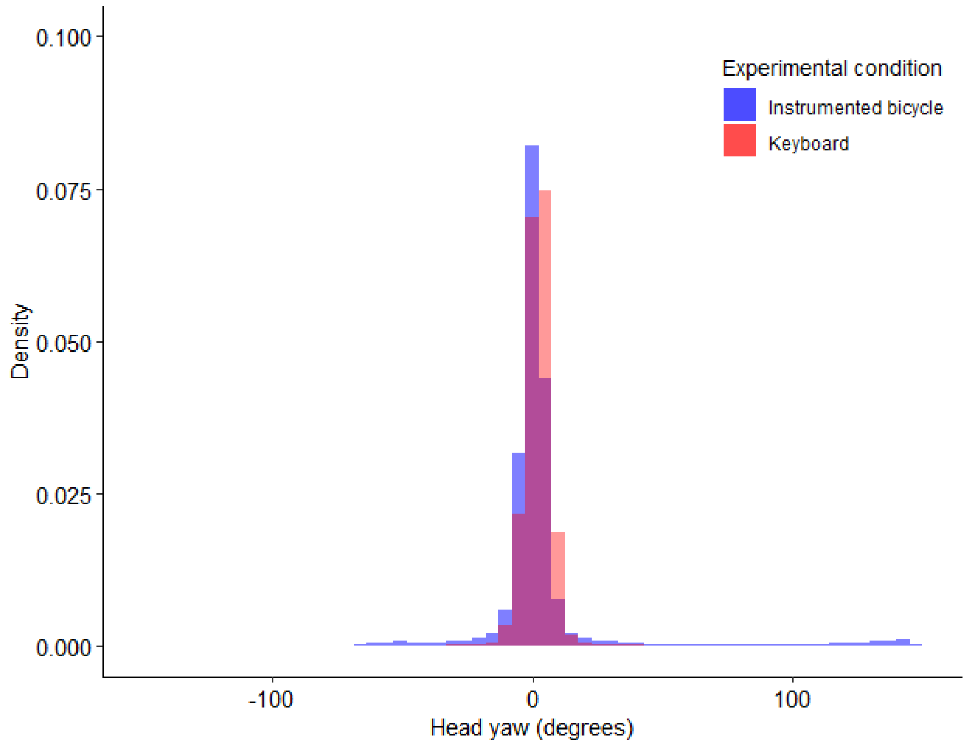

Figure 5: Histograms of Head Yaw in the two experimental settings.

Acceleration

The results of the one-sided F-test showed that the variance of Acceleration is significantly larger in the instrumented-bike experiment (F (47,45) = 3.1474, p-value = 0.00008789). This is in line with our Hypothesis 3: we expected higher variation in acceleration due to the presence of physical effort and higher difficulty in maintaining constant speed. The density histogram for Acceleration presented in **Figure 6** reflects the results of the test, as we observe that the Acceleration values are accumulated near the mean with little variation in the keyboard-controlled experiment and have a near bell-shape distribution in the other experiment.

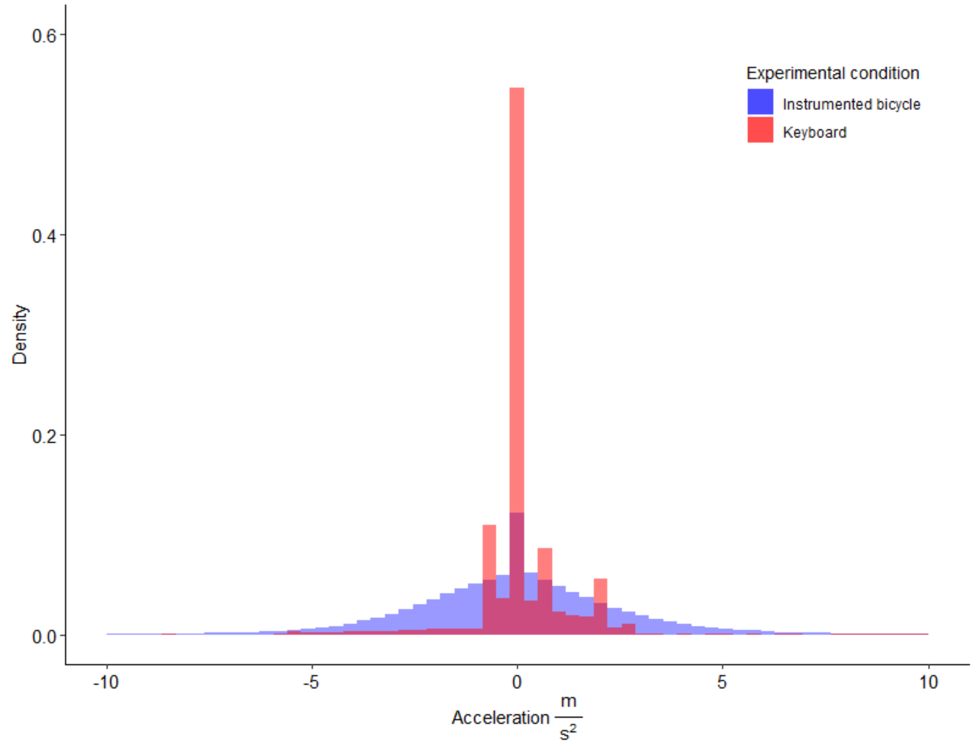

Figure 6: Histograms of Acceleration in the two experimental settings.

Braking

The result of one-sided F-test showed that the variance of Braking is significantly larger in the instrumented-bike experiment (F (47,45) = 3.8141, p-value = 0.000007114). This result conforms to our Hypothesis 4 that more variance is present in the instrumented-bicycle experiment. Again, the density histogram of Braking in **Figure 7** visually reflects the results of the test, as there is more variation in an experiment that used an instrumented bicycle.

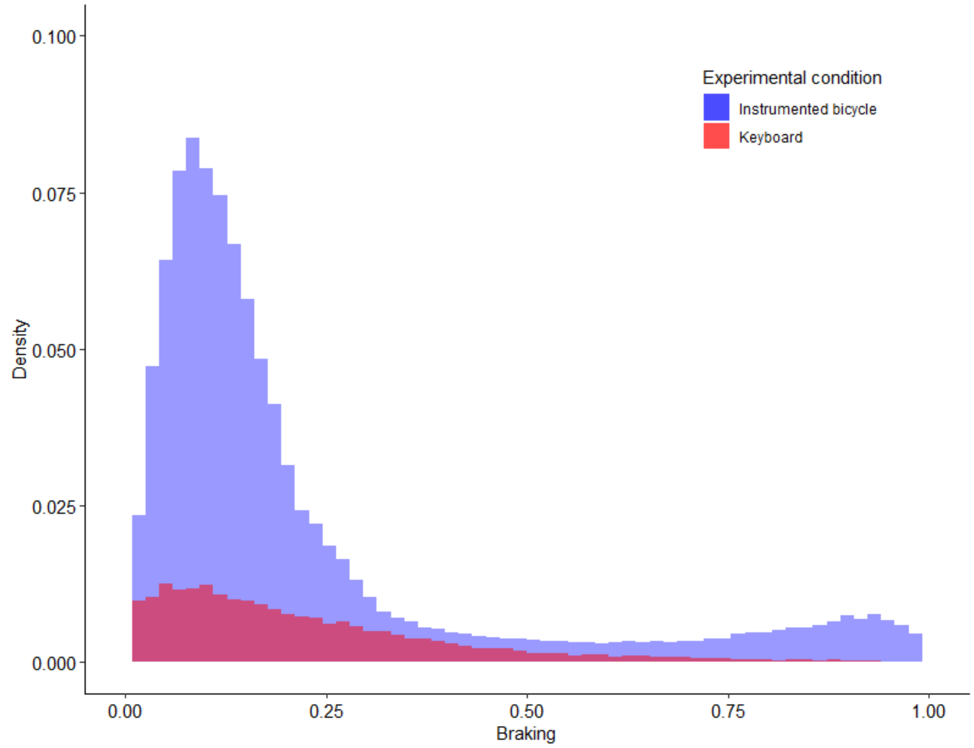

Figure 7: Histograms of Braking in the two experimental settings.

Amplitude of α wave

The one-sided Welch's t-test demonstrated that the mean α amplitude is lower in the instrumented bicycle experiment as opposed to the keyboard-controlled at the 90% confidence level (t = 1.35, df = 40.79, p-value = 0.09), which is in line with our Hypothesis 5, where we expected a lower α amplitude in instrumented bicycle experiment due to the more immersive setting and higher cognitive engagement compared to the keyboard.

DISCUSSION

VR experiments can effectively contribute to transport research only if the behaviour elicited in a virtual environment closely resembles real-world behaviour. Hence, it is important to be able to discriminate between different experimental designs that employ immersive technologies. The objective of the present paper was to compare the cycling behaviour elicited in two separate experiments which used the same visual stimuli but different devices to control the navigation through the simulated scenarios. The first one employed a keyboard and the second one an instrumented bicycle. In order to draw conclusions about the behaviour, we analysed participants' speed, acceleration, braking and head movements, along with data about their neural activity. The results are summarized in **Table 2**, where we also show examples from the existing literature which reached similar results in different experimental contexts. This highlights not only that our findings are in line with the literature, but that we confirm these results for the context of cycling and for the joint use of immersive technology and EEG.

Table 2: Summary of results

No.	Hypotheses	Result	Similar conclusions in the literature

1A	Average speed is higher in the keyboard-controlled experiment than in instrumented bicycle one.	True	Bella (51) with driving simulator
1B	There is more variance in speed in the instrumented bicycle experiment than in the keyboard experiment.	True	Godley et al. (8) with driving simulator
2	The average head movement is higher in the instrumented bicycle experiment than in the keyboard experiment.	True	Sitzmann et al. (26) focusing on head movement in VR; Underwood et al. (7) with driving simulator context
3	There is more variance in the acceleration behaviour in the instrumented bicycle experiment than in the keyboard experiment.	True	Reymond et al. (52) with driving simulator
4	There is more variance in the braking behaviour in the instrumented bicycle experiment than in the keyboard experiment.	True	Godley et al. (8) focusing on braking behaviour in driving simulator
5	The mean amplitude of α wave is higher in the keyboard-controlled experiment compared to an instrumented bicycle.	True	Argento et al. (53) in the context of brainwave entertainment in VR

Overall, the results show significant differences in behaviour. A more active and varied behaviour is observed in the instrumented bicycle experiment as compared to the keyboard-controlled study. For instance, the average speed is lower and more heterogeneous in the instrumented bicycle experiment, suggesting that participants dedicate more time to explore the environment. Moreover, the average speed in the instrumented bicycle experiment of 14.32 km/h is closer to the average speed of cycling in reality, which ranges between 13.5-16 km/h, in contrast to the keyboard counterpart where average speed was considerably higher in spite of the cap of 25km/h. It suggests that the use of the instrumented bicycle instead of the keyboard allows for better approximation of the real cycling kinematics.

More variation in acceleration and braking, as well as more head movement is also observed in the instrumented bicycle setting, implying a higher degree of engagement with the environment. This is further confirmed by the analysis of the EEG data, where a lower amplitude of the α wave in the instrumented bicycle experiment suggests higher mental engagement in the task compared to the keyboard-controlled one.

Our work provides evidence that the instrumented bicycle is more effective than the keyboard controls in eliciting behavioural patterns demonstrated by previous naturalistic studies of cycling behaviour. We use these studies as a benchmark due to the lack of evidence in the previous literature with regard to typical, real-life cycling behaviour in absolute terms. Moreover, our results are consistent with previous studies conducted in other contexts that investigated the effects of the use of various input devices on behaviour, as presented in the introduction and in **Table 2**. The use of a keyboard, or in fact any other input device such as a joystick or touchpad, makes the cycling experience clearly less realistic as the user does not need to exert physical effort, a crucial component associated with cycling. It then follows that changes in action could be seen as less consequential as they do not impact on physical fatigue. This relates both to the actual action (e.g. accelerate) as well as the degree thereof. On the other hand, the advantages of the use of a keyboard or other simple devices cannot be ignored. Their use as an input device is cheaper and less time consuming than the employment of the stationary bicycle. They are also more portable which makes it easier to collect data in different locations. Moreover, if VR is used jointly with neuroimaging devices, then a simpler input device offers a more static approach, reducing the extent of potential noise in the neural data stemming from body movements. The choice of the input device for VR experiments is thus not an arbitrary decision and should be aligned with the objectives of the study to not constrain the spectrum of behaviour which can be captured and minimise the potential biases resulting from the mere choice of the controller. The decision of an appropriate input tool also has to be weighed against technical capabilities such the budget, duration of the experiment and comfort of the participants as well as a possibility of joint use with other equipment employed in the study. In this paper, we compare only two devices. However, it is important to take into consideration other available appliances such as a 3D mouse, joystick, steering wheel, gamepad or hybrid controller which may offer different benefits depending on the design of the study.

The results thus emphasize the importance of the experimental setup in a VR experiment beyond the choice of appropriate visual stimulus. The findings extend understanding of the effects of the use of distinct input devices within the VR domain by demonstrating that the use of an instrumented bicycle increases the realism of the cycling simulations by influencing the manoeuvring decisions. These results were further reinforced by the analysis of neural data. Further research needs to be conducted to generalize these findings. In particular, we recommend testing different cycling scenarios as well as experiments focusing on different aspects of travel behaviour to compare participants' actions in the experimental setting with real-world decisions. The findings shed light on the level of behavioural congruence of the state-of-the-art VR studies which will be valuable in the interpretation and the level of confidence in the results of different VR studies. It is also expected to be a valuable resource to researchers and practitioners planning to administer VR-based data collection and help them to better design the experimental setup as there has been a significant interest in using VR for modelling cycling behaviour (54-55). By comparing and contrasting the behaviour of cyclists in the two VR environments, the paper is expected to provide guidance to researchers investigating cycling behaviour in dynamic setting and hence improve the modelling of speed and acceleration of cyclists which can feed into safety research or capacity analyses for instance. The findings are also expected to be useful for planners who are interested in deploying VR to more realistically test the impact of different urban designs on the propensity to cycle, indicating, for example, the road and pavements features which contribute to the higher perception of safety among cyclists. The research findings can hence help transport and urban planners in making more informed choices regarding urban infrastructure. Finally, VR tools are increasingly being used in designing vehicles of the future – the interaction between connected and automated vehicles (CAVs) and other road users. The findings can help researchers modelling the interaction between cyclists and CAVs in designing their experiments and better interpreting the results by giving an idea about the comparative realism of the collected data.

ACKNOWLEDGMENTS

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

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception and design: M. Bogacz, S. Hess, C. Choudhury, C. Calastri, A. Erath, M. Van Eggermond ; data collection: M. Bogacz, M. Nazemi; analysis and interpretation of results: M. Bogacz, S. Hess, C. Choudhury, C. Calastri, F. Mushtaq; draft manuscript preparation: M. Bogacz, S. Hess, C. Choudhury, C. Calastri. All authors reviewed the results and approved the final version of the manuscript.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

1. Mai, K. L. (2017). Evaluation of PC-Based Virtual Reality as a Tool to Analyze Pedestrian Behavior at Midblock Crossings. (Master's thesis, California Polytechnic State University).

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

3. Moussa, G., Radwan, E., & Hussain, K. (2012). Augmented reality vehicle system: Left-turn manoeuvre study. Transportation research part C: emerging technologies, 21(1), 1-16.

4. Farooq, B. Cherchi, E., & Sobhani, A. (2018). Virtual immersive reality for travel behaviour analysis: A case study of autonomous vehicles on urban roads. forthcoming Journal of the Transportation Research Board.

5. Bogacz, M., Calastri, C., Choudhury, C. F., Hess, S., Erath, A., van Eggermond, M. A., & Mushtag, F. (2018, August). Processing cycling risk under different elicitation methods: comparing 2D and 3D in virtual reality choice environments. Presented at 99th Annual Meeting of the Transportation Research Board, Washington, D.C., 2019.

6. Patterson, Z., Mostofi-Darbani, J., Rezaei, A. & Zacharias, J. (2017) Comparing Text-only and Virtual Reality Discrete Choice Experiments of Neighborhood Choice. Landscape and Urban Planning, 157, 63-74.

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

8. Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. Accident analysis & prevention, 34(5), 589-600.

9. Szul, M. J., Bompas, A., Sumner, P., & Zhang, J. (2019). The validity and consistency of continuous joystick response in perceptual decision-making. bioRxiv, 501536.

10. Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron, 45, 801–814.

11. Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. Proceedings of the National Academy of Sciences, 102, 10393–10398.

12. Rupp, M. A., Oppold, P., & McConnell, D. S. (2015). Evaluating input device usability as a function of task difficulty in a tracking task. Ergonomics, 58(5), 722-735.

13. Chung, S., Kramer, T., & Wong, E. M. (2018). Do touch interface users feel more engaged? The impact of input device type on online shoppers' engagement, affect, and purchase decisions. Psychology & Marketing, 35(11), 795-806.

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

15. Lal, S. K., & Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. Biological psychology, 55(3), 173-194.

16. Eoh, H. J., Chung, M. K., & Kim, S. H. (2005). Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. International Journal of Industrial Ergonomics, 35(4), 307-320.

17. Craig, A., Tran, Y., Wijesuriya, N., & Nguyen, H. (2012). Regional brain wave activity changes associated with fatigue. Psychophysiology, 49(4), 574-582.

18. Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain research reviews, 29(2-3), 169-195.

19. Puce, A., & Hämäläinen, M. (2017). A review of issues related to data acquisition and analysis in EEG/MEG studies. Brain sciences, 7(6), 58.

20. Duvinage, M., Castermans, T., Dutoit, T., Petieau, M., Hoellinger, T., De Saedeleer, C., ... & Cheron, G. (2012). A P300-based quantitative comparison between the Emotiv Epoc headset and a medical EEG device. Biomedical Engineering, 765(1), 2012-2764.

21. Ratti, E., Waninger, S., Berka, C., Ruffini, G., & Verma, A. (2017). Comparison of medical and consumer wireless EEG systems for use in clinical trials. Frontiers in human neuroscience, 11, 398.

22. Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., & Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. Neuroscience letters, 244(2), 73-76.

23. Sauseng, P., Klimesch, W., Stadler, W., Schabus, M., Doppelmayr, M., Hanslmayr, S., ... & Birbaumer, N. (2005). A shift of visual spatial attention is selectively associated with human EEG alpha activity. European Journal of Neuroscience, 22(11), 2917-2926.

24. Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). α-Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. Journal of Neuroscience, 26(37), 9494-9502.

25. Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. Trends in cognitive sciences, 16(12), 606-617.

26. Sitzmann, V., Serrano, A., Pavel, A., Agrawala, M., Gutierrez, D., Masia, B., & Wetzstein, G. (2016). How do people explore virtual environments?. arXiv preprint arXiv:1612.04335.

27. Hu, B., Johnson-Bey, I., Sharma, M., & Niebur, E. (2017, March). Head movements during visual exploration of natural images in virtual reality. In 2017 51st Annual Conference on Information Sciences and Systems (CISS) (pp. 1-6). IEEE.

28. Hawkins, G. E., Mittner, M., Boekel, W., Heathcote, A., & Forstmann, B. U. (2015). Toward a model-based cognitive neuroscience of mind wandering. Neuroscience, 310, 290-305.

29. Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Frontiers in human neuroscience, 4, 186.

30. Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. Brain research reviews, 53(1), 63-88.

31. EMOTIV EPOC+ - 14 Channel Wireless EEG Headset. (2018). Available at: https://www.emotiv.com/epoc/. [Accessed: 17.12.2018]

32. Oculus. (2018). Available at: https://www.oculus.com. [Accessed: 17.12.2018]

33. Schramka, F., Arisona, S., Joos, M., & Erath, A. (2017). Development of Virtual Reality Cycling Simulator, in 3rd International Conference on Virtual Reality, paper presented at 3rd International Conference on Virtual Reality, Hong Kong.

34. Unity. (2017). Unity - Game Engine. [online] Available at: https://unity3d.com [Accessed 10.11. 2018].

35. Vive.com. (2019). *VIVE™ | Discover Virtual Reality Beyond Imagination*. [online] Available at: https://www.vive.com/uk/ [Accessed 2.08.2019].

36. Neuroelectrics. (2019). Products / ENOBIO / ENOBIO 8 5G. [online] Available at: https://www.neuroelectrics.com/products/enobio/enobio-8-5g/ [Accessed 2.08.2019].

37. Di Stasi, L. L., Renner, R., Catena, A., Cañas, J. J., Velichkovsky, B. M., & Pannasch, S. (2012). Towards a driver fatigue test based on the saccadic main sequence: A partial validation by subjective report data. Transportation research part C: emerging technologies, 21(1), 122-133.

38. Katsis, C. D., Goletsis, Y., Rigas, G., & Fotiadis, D. I. (2011). A wearable system for the affective monitoring of car racing drivers during simulated conditions. Transportation research part C: emerging technologies, 19(3), 541-551.

39. Welch, B. L. (1947). The generalization of student's' problem when several different population variances are involved. Biometrika, 34(1/2), 28-35.

40. Johnston, J. (1972). Econometric Methods (Second ed.). New York: McGraw-Hill.

41. Dozza, M., & Werneke, J. (2014). Introducing naturalistic cycling data: What factors influence bicyclists' safety in the real world?. Transportation research part F: traffic psychology and behaviour, 24, 83-91.

42. Schleinitz, K., Petzoldt, T., Franke-Bartholdt, L., Krems, J., & 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.

Bogacz, M., Hess, S., Calastri, C., Choudhury, C. F., Erath, A., van Eggermond, M. A., Mushtaq, F., Nazemi, M. & Awais, M.

43. Huertas-Leyva, P., Dozza, M., & Baldanzini, N. (2018). Investigating cycling kinematics and braking maneuvers in the real world: e-bikes make cyclists move faster, brake harder, and experience new conflicts. Transportation research part F: traffic psychology and behaviour, 54, 211-222.

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

45. Repovs, G. (2010). Dealing with noise in EEG recording and data analysis. In Informatica Medica Slovenica, 15(1), 18-25.

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

47. Ille, N., Berg, P., & 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.

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

49. Shaker, M. M. (2006). EEG waves classifier using wavelet transform and Fourier transform. brain, 2, 3.

50. Oken, B. S., & Chiappa, K. H. (1986). Statistical issues concerning computerized analysis of brainwave topography. Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society, 19(5), 493-494.

51. Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. Accident Analysis & Prevention, 40(3), 1078-1087.

52. Reymond, G., Kemeny, A., Droulez, J., & Berthoz, A. (2001). Role of lateral acceleration in curve driving: Driver model and experiments on a real vehicle and a driving simulator. Human factors, 43(3), 483-495.

53. Argento, E., Papagiannakis, G., Baka, E., Maniadakis, M., Trahanias, P., Sfakianakis, M., & Nestoros, I. (2017). Augmented Cognition via Brainwave Entrainment in Virtual Reality: An Open, Integrated Brain Augmentation in a Neuroscience System Approach. Augmented Human Research, 2(1), 3.

54. Nazemi, M., van Eggermond, M. A., Erath, A., & Axhausen, K. W. (2018). Studying cyclists' behavior in a non-naturalistic experiment utilizing cycling simulator with immersive virtual reality. Arbeitsberichte Verkehrs-und Raumplanung, 1383.

55. Nazemi, M., van Eggermond, M. A., Erath, A., Schaffner, D., Joos, M., & Axhausen, K. W. (2019). Studying bicyclists' perceived level of safety using a cycling simulator combined with immersive virtual reality. In 8th International Cycling Safety Conference (ICSC 2019).